(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 January 2005 (13.01.2005)

PCT

(10) International Publication Number
WO 2005/003778 A2

(51) International Patent Classification[7]: G01N 33/68, C12Q 1/68, C07B 61/00

(21) International Application Number: PCT/DK2004/000466

(22) International Filing Date: 30 June 2004 (30.06.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
PA 2003 01003 2 July 2003 (02.07.2003) DK
60/483,899 2 July 2003 (02.07.2003) US

(71) Applicant *(for all designated States except US)*: **NUEVOLUTION A/S** [DK/DK]; Rønnegade 8, 5., DK-2100 Copenhagen Ø (DK).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **FRESKGÅRD, Per-Ola** [SE/SE]; Ringvägen 7, S-235 93 Vellinge (SE). **LUNDORF, Mikkel, Dybro** [DK/DK]; Viborggade 13, st.tv., DK-2100 København O (DK). **JAKOBSEN, Søren, Nyboe** [DK/DK]; Frederiksvej 36, 4.tv., DK-2000 Frederiksberg (DK).

(74) Agent: **HØIBERG A/S**; St. Kongensgade 59A, DK-1264 Copenhagen K (DK).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: A METHOD FOR IDENTIFYING A SYNTHETIC MOLECULE HAVING AFFINITY TOWARDS A TARGET

(57) Abstract: The present invention relates to a method for identifying a synthetic molecule having affinity towards a target. The method includes the steps of providing a library of bifunctional complexes, wherein each complex of the library comprises a synthetic molecule attached to an identifier, which codes for said molecule; subjecting, under binding conditions, the library of bifunctional complexes to a target; removing the non-binding members of the library; separating the identifiers of complexes comprising synthetic molecules having affinity towards the target, and decoding the identifiers to establish the identity of the molecule.

WO 2005/003778 A2

A METHOD FOR IDENTIFYING A SYNTHETIC MOLECULE HAVING AFFINITY TOWARDS A TARGET

This application is a non-provisional of U.S. provisional application Serial No. 60/483,899 filed on 2 July 2003, which is hereby incorporated by reference in its entirety. All patent and non-patent references cited in the application, or in the present application, are also hereby incorporated by reference in their entirety.

Technical Field of the Invention

The present invention relates to a method for identifying from a library a synthetic molecule having affinity towards a target. The synthetic molecule has initially been a part of a complex also comprising an identifier that codes for said synthetic molecule. The invention further relates to a library of complexes, which in a certain embodiment can be used in the method. The experiments supporting the present method suggest an enrichment of more than a million times, e.g. a specific synthetic molecule may be identified in a library of 10^6+ complexes.

Background

Traditional drug discovery begins with a pathological phenomenon in an organism and the development of a therapeutic theory to combat this. A chemical concept follows to produce compounds for screening. Most of the processes for curing the pathological phenomenon originate with the understanding of some biological pathways and screening for an effect in tissues or cells. This may or may not eventually reveal a "target". The target can be identified by various conventional methods, including protein expressing, protein chemistry, structure-functional studies, knowledge of biochemical pathways, and genetic studies.

In recent years, genetic information has increasingly guided the identification of single molecular targets. These are derived from the knowledge of the genes of specific cell phenotypes that encode proteins that may be involved in the pathogenesis of a particular disease state.

A lead is a compound, usually a small organic molecule that demonstrates a desired biological activity on a target. Usually, a collection of compounds, referred to as a "library", is screened before a useful lead is identified. Today, many libraries are

commercially available or open to public. Most pharmaceutical companies house their own compilation of compounds that have been synthesised over several years and screened against a variety of targets.

Each compound in a library must be screened by an appropriate assay against the target. Techniques for handling the screening of several thousands compounds simultaneously have been developed and are generally referred to as high-throughput screening techniques. To push the limit of compounds possible to screen simultaneously, different manufactures have been developing instrumentation capable of handling multiple micro titer plate formats on the same platform using 384 and 1536-well plates. Advances in small volume liquid dispensing and pipetting, reliable handling of standardized plates and simplified assay formats all have made an impact on the reliability of the high-throughput screening process.

However, high-throughput screening has the disadvantage that each of the compounds have to be positioned in spatially discrete regions, usually in wells of a micro titer plate in order to observe an interaction with a target. If more than a single compound is present, it is not feasible to discern which compound displaying the appropriate biological activity. Thus, the full power of combinatorial chemistry cannot be applied because a collection of compounds usually is produced in a single container.

To be able to select a possible lead compound in a collection of compounds placed in the same contained, libraries of bifunctional complexes have been evolved. Each bifunctional complex in the library comprises a potential lead compound coupled to an identifier sequence. The identifier sequence is suitably a nucleic acid which uniquely identifies the potential lead compound. When a library of bifunctional complexes is screened against a target, one or more of the potential leads may bind to the target. After removal of the remainder of the library, the binding bifunctional complexes can be eluated and the lead compound identified by sequencing the identifier.

Various techniques for producing bifunctional complexes are known from the prior art. Some attempts to form the complex comprising a molecule as well as the identifier that codes therefore, are based on the split-and-mix principle known from combinatorial chemistry, see e.g. WO 93/06121 A1, EP 643 778 B1, and WO 00/23458. Other attempts have focussed on the formation of encoded proteins using the natural machinery of a cell and connecting the formed protein with the template nucleic acid that has coded for the amino acid components of the protein. Examples of suitable systems are phage display, *E. coli* display, ribosome display (WO 93/03172), and protein-mRNA-fusions (WO 98/31700). The genetic information of the nucleic acid, usually mRNA or DNA, may not necessarily be decoded between each round of selection to establish the identity of the chemical entities that has formed the protein because the nucleic acid can be amplified by known means, such as PCR, and processed for the formation of a new library enriched in respect of suitable binding proteins.

Recently, a method for encoding molecules has been suggested, which can be performed in several selection rounds without intermediate decoding, wherein the encoded molecule is not restricted to peptides and proteins. WO 02/00419 and WO 02/103008 disclose methods for preparing virtually any molecule connected to an identifier coding for chemical entities which have reacted to form the molecule. In short, a template segregated into a plurality of codons and a plurality of building blocks comprising a transferable chemical entity and an anticodon are initially provided. Under hybridisation conditions, the template and building blocks are annealed together and the chemical entities are subsequently reacted to form the molecule.

The present invention aims at providing an efficient method for identifying molecules having affinity towards a target using a library of bifunctional complexes.

Summary of the Invention

The present invention concerns a method for identifying a synthetic molecule having affinity towards a target, comprising the steps of

a) providing a library of bifunctional complexes, wherein each complex of the library comprises a synthetic molecule attached to an identifier, which codes for said molecule,

b) subjecting, under binding conditions, the library of bifunctional complexes to a target,

c) removing the non-binding members of the library,

d) separating the identifiers of complexes comprising synthetic molecules having affinity towards the target, and

e) decoding the identifiers to establish the identity of the molecule.

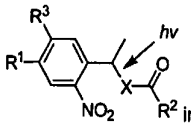
The present invention offers a novel method for enrichment of libraries of bifunctional complexes. In certain aspects, the method takes advantage of retaining the non-specific binding complexes bound to the target, walls of wells, beads etc while the identifiers of the complexes having specific affinity towards the target are separated. Thus, the present invention provides in certain aspects a method not only selecting such complexes which survives a selection assay, as the non-specific binding complexes are excluded from the pool of identifier separated.

While it is not desired to be bound by any particular theory, it is presently believed that the identifier part of the complexes possesses an inherent affinity towards the surfaces of a reaction chamber or parts of the target not involved in the binding of an agonist or antagonist. The separation step ensures that the non-specific binding complexes are retained in the vessel. The identifiers attached to the synthetic molecule may be separated by cleaving different kind of linkages. Preferably the linkage is selectively cleavable, i.e. when the media comprising the library of complexes is exposed to a certain condition, only the intended linkage is cleaved.

In one embodiment of the invention, a cleavable linker moiety is situated between the synthetic molecule and the identifier. When the complex is subjected to conditions cleaving the linker, only such identifiers that are not adhered to solid surfaces etc. are liberated into the liquid media. The identifiers in the liquid media may easily be recovered, thereby separating these from the non-specific binding identifiers remaining immobilized. In another embodiment of the invention, a cleavable linker is positioned between a target and a solid surface, thereby immobilizing the target. Subsequent to exposing the immobilized target to a library of complexes and removing the non-binding part of the library, the complexes attached to the target may separated by exposing the media to a condition cleaving the linkage between the target and the solid surface. If desirable the identifier can be isolated and amplified using standard molecular methods, such as PCR.

The second embodiment may be combined with chromatography. Following the separation of the complexes attached to the target, the media may be subjected to chromatography, e.g. size exclusion chromatography. In certain embodiments after the chromatography process another orthogonal cleavable linkage between the synthetic molecule and the identifier may be cleaved to liberate the identifier.

The invention also relates to a library of complexes, in which each different complex comprises a synthetic molecule, attached via a cleavable linkage to an identifier which codes for said molecule. The linkage may be cleaved by a variety of different conditions, such as electromagnetic radiation, chemical agents, and enzymes. In a preferred aspect the selective cleavable linkage comprises a chemical moiety, which can be cleaved by electromagnetic irradiation, such as light having a specified wave length. A useful chemical moiety for the cleavable linkage comprises a group



in which R^1 and R^2 are either of the synthetic molecule or the identifier, respectively, R^3 is H or OCH_3, and X is O, S, or NH. The identifiers preferably comprise a sequence of nucleotides. In one aspect of the invention, the identifier comprises 2 or more codons, which codes for 2 or more chemical entities incorporated into the synthetic molecule.

Complex

The complex of the present invention comprises a synthetic molecule and an identifier. The identifier comprises identifying moieties that identifies the molecule. Preferably, the identifier identifies the molecule uniquely, i.e. in a library of complexes a particular identifier is capable of distinguishing the molecule it is attached to from the rest of the molecules.

The molecule and the identifier may be attached directly to each other or through a bridging moiety. In one aspect of the invention, the bridging moiety is a selectively cleavable linkage. The identifying moieties of each complex suitably comprise recognition units, i.e. units which may be recognized by a detecting entity. A variety of different kinds of recognition exist in nature. Examples include antibodies which recognize an epitope, proteins which recognize another protein, mRNA which recognize a protein, small molecules (like biotin) which recognize a protein (like avidine or

streptavidine) and oligonucleotides which recognizes complementing oligonucleotides. Generally it is preferred that the identifier is a sequence of nucleotides.

The method may in certain embodiments be performed without amplification after the separation step. However, when larger libraries are used and the amount of separated identifiers is relatively low, it is in general preferred to use an identifier which is amplifiable. Identifiers comprising a sequence of nucleotides may be amplified using standard techniques, like PCR. In the event the identifier is a protein, the protein may be amplified by attaching the mRNA which encoded the synthesis thereof, generating the cDNA from the mRNA and subjecting said mRNA to a translation system. Such system is described in WO 98/31700 the content of which is incorporated herein by reference. An alternative method for amplifying a protein is to use phage-displayed proteins.

The identifier may comprise two or more codons. The sequence of codons can be decoded to identify reactants used in the formation of the molecule. When the identifier comprises more than one codon, each member of a pool of building blocks can be identified uniquely and the order of codons is informative of the synthesis step each member has been incorporated in.

The sequence of the nucleotides in each codon may have any suitable length. The codon may be a single nucleotide or a plurality of nucleotides. In some aspects of the invention, it is preferred that each codon independently comprises four or more nucleotides, more preferred 4 to 30 nucleotides.

The identifier will in general have at least two codons arranged in sequence, i.e. next to each other. Two neighbouring codons may be separated by a framing sequence. Depending on the encoded molecule formed, the identifier may comprise further codons, such as 3, 4, 5, or more codons. Each of the further codons may be separated by a suitable framing sequence. Preferably, all or at least a majority of the codons of the identifier are separated from a neighbouring codon by a framing sequence. The framing sequence may have any suitable number of nucleotides, e.g. 1 to 20. Alternatively, codons on the identifier may be designed with overlapping sequences.

The framing sequence, if present, may serve various purposes. In one setup of the invention, the framing sequence identifies the position of the codon. Usually, the framing sequence either upstream or downstream of a codon comprises information which allows determination of the position of the codons. In another setup, the frames have alternating sequences, allowing for addition of building blocks from two pools in the formation of the library. The framing sequence may also or in addition provide for a region of high affinity. The high affinity region may ensure that the hybridisation of the template with an anti-codon will occur in frame. Moreover, the framing sequence may adjust the annealing temperature to a desired level.

A framing sequence with high affinity can be provided by incorporation of one or more nucleobases forming three hydrogen bonds to a cognate nucleobase. Examples of nucleobases having this property are guanine and cytosine. Alternatively, or in addition, the framing sequence may be subjected to backbone modification. Several back bone modifications provides for higher affinity, such as 2'-O-methyl substitution of the ribose moiety, peptide nucleic acids (PNA), and 2'-4' O-methylene cyclisation of the ribose moiety, also referred to as LNA (Locked Nucleic Acid).

The identifier may comprise flanking regions around the codons. The flanking region can encompass a signal group, such as a flourophor or a radio active group to allow for detection of the presence or absence of a complex or the flanking region may comprise a label that may be detected, such as biotin. When the identifier comprises a biotin moiety, the identifier may easily be recovered following the separation step.

The flanking regions can also serve as priming sites for amplification reactions, such as PCR. The identifier may in certain embodiments comprise an affinity region having the property of being able to hybridise to a building block.

It is to be understood that when the term identifier is used in the present description and claims, the identifier may be in the sense or the anti-sense format, i.e. the identifier can be a sequence of codons which actually codes for the molecule or can be a sequence complementary thereto. Moreover, the identifier may be single-stranded or double-stranded, as appropriate.

The synthetic molecule part of the complex is generally of a structure expected of having an effect on the target. When the target is of pharmaceutical importance, the molecule is generally a possible drug candidate. The complex may be formed by tagging a library of different possible drug candidates with a tag, e.g. a nucleic acid tag uniquely identifying each possible drug candidate. In another embodiment of the invention, the molecule is encoded, i.e. formed by a variety of reactants which have reacted with each other and/or a scaffold molecule. Optionally, this reaction product may be post-modified to obtain the final molecule displayed on the complex. The post-modification may involve the cleavage of one or more chemical bonds attaching the encoded molecule to the indentifier in order more efficiently to display the encoded molecule.

The formation of an encoded molecule generally starts by a scaffold, i.e. a chemical unit having one or more reactive groups capable of forming a connection to another reactive group positioned on a chemical entity, thereby generating an addition to the original scaffold. A second chemical entity may react with a reactive group also appearing on the original scaffold or a reactive group incorporated by the first chemical entity. Further chemical entities may be involved in the formation of the final reaction product. The formation of a connection between the chemical entity and the nascent encoded molecule may be mediated by a bridging molecule. As an example, if the nascent encoded molecule and the chemical entity both comprise an amine group a connection between these can be mediated by a dicarboxylic acid. A synthetic molecule is in general produced in vitro and may be a naturally occurring or an artificial substance. Usually, a synthetic molecule is not produced using the naturally translation system in an in vitro process.

The chemical entities that are precursors for structural additions or eliminations of the encoded molecule may be attached to a building block prior to the participation in the formation of the reaction product leading the final encoded molecule. Besides the chemical entity, the building block generally comprises an anti-codon. In some embodiments the building blocks also comprise an affinity region providing for affinity towards the nascent complex.

Thus, the chemical entities are suitably mediated to the nascent encoded molecule by a building block, which further comprises an anticodon. The anti-codon serves

the function of transferring the genetic information of the building block in conjunction with the transfer of a chemical entity. The transfer of genetic information and chemical entity may occur in any order, however, it is important that a correspondence is maintained in the complex. The chemical entities are preferably reacted without enzymatic interaction in some aspects of the invention. Notably, the reaction of the chemical entities is preferably not mediated by ribosomes or enzymes having similar activity. In another aspect of the invention a ribosome is used to translate an mRNA into a protein using a tRNA loaded with a natural or unnatural amino acid.

According to certain aspects of the invention the genetic information of the anti-codon is transferred by specific hybridisation to a codon on a nucleic acid template. Another method for transferring the genetic information of the anti-codon to the nascent complex is to anneal an oligonucleotide complementary to the anti-codon and attach this oligonucleotide to the complex, e.g. by ligation. A still further method involves transferring the genetic information of the anti-codon to the nascent complex by an extension reaction using a polymerase and a mixture of dNTPs.

The chemical entity of the building block may in most cases be regarded as a precursor for the structural entity eventually incorporated into the encoded molecule. In other cases the chemical entity provides for the eliminations of chemical units of the nascent encoded molecule. Therefore, when it in the present application with claims is stated that a chemical entity is transferred to a nascent encoded molecule it is to be understood that not necessarily all the atoms of the original chemical entity is to be found in the eventually formed encoded molecule. Also, as a consequence of the reactions involved in the connection, the structure of the chemical entity can be changed when it appears on the nascent encoded molecule. Especially, the cleavage resulting in the release of the entity may generate a reactive group which in a subsequent step can participate in the formation of a connection between a nascent complex and a chemical entity.

The chemical entity of the building block comprises at least one reactive group capable of participating in a reaction which results in a connection between the chemical entity of the building block and another chemical entity or a scaffold associated with the nascent complex. The number of reactive groups which appear on the

chemical entity is suitably one to ten. A building block featuring only one reactive group is used i.a. in the end positions of polymers or scaffolds, whereas building blocks having two reactive groups are suitable for the formation of the body part of a polymer or scaffolds capable of being reacted further. One, two or more reactive groups intended for the formation of connections, are typically present on scaffolds. Non-limiting examples of scaffolds are opiates, steroids, benzodiazepines, hydantoines, and peptidylphosphonates.

The reactive group of the chemical entity may be capable of forming a direct connection to a reactive group of the nascent complex or the reactive group of the building block may be capable of forming a connection to a reactive group of the nascent complex through a bridging fill-in group. It is to be understood that not all the atoms of a reactive group are necessarily maintained in the connection formed. Rather, the reactive groups are to be regarded as precursors for the structure of the connection.

The subsequent cleavage step to release the chemical entity from the building block can be performed in any appropriate way. In an aspect of the invention the cleavage involves usage of a chemical reagent or an enzyme. The cleavage results in a transfer of the chemical entity to the nascent encoded molecule or in a transfer of the nascent encoded molecule to the chemical entity of the building block. In some cases it may be advantageous to introduce new chemical groups as a consequence of linker cleavage. The new chemical groups may be used for further reaction in a subsequent cycle, either directly or after having been activated. In other cases it is desirable that no trace of the linker remains after the cleavage.

In another aspect, the connection and the cleavage is conducted as a simultaneous reaction, i.e. either the chemical entity of the building block or the nascent encoded molecule is a leaving group of the reaction. In some aspects of the invention, it is appropriate to design the system such that the connection and the cleavage occur simultaneously because this will reduce the number of steps and the complexity. The simultaneous connection and cleavage can also be designed such that either no trace of the linker remains or such that a new chemical group for further reaction is introduced, as described above.

The attachment of the chemical entity to the building block, optionally via a suitable spacer can be at any entity available for attachment, e.g. the chemical entity can be attached to a nucleobase or the backbone. In general, it is preferred to attach the chemical entity at the phosphor of the internucleoside linkage or at the nucleobase. When the nucleobase is used for attachment of the chemical entity, the attachment point is usually at the 7 position of the purines or 7-deaza-purins or at the 5 position of pyrimidines. The nucleotide may be distanced from the reactive group of the chemical entity by a spacer moiety. The spacer may be designed such that the conformational spaced sampled by the reactive group is optimized for a reaction with the reactive group of the nascent encoded molecule.

The synthetic molecules of the invention may have any chemical structure. In a preferred aspect, the synthetic molecule can be any compound that may be synthesized in a component-by-component fashion. In some aspects the synthetic molecule is a linear or branched polymer. In another aspect the synthetic molecule is a scaffolded molecule. The term "synthetic molecule" also comprises naturally occurring molecules like α-polypeptides etc, however produced *in vitro* usually in the absence of enzymes, like ribosomes. In certain aspects, the synthetic molecule of the library is a non-α-polypeptide.

The synthetic molecule may have any molecular weight. However, in order to be orally available, it is in this case preferred that the synthetic molecule has a molecular weight less than 2000 Daltons, preferably less than 1000 Dalton, and more preferred less than 500 Daltons.

The size of the library may vary considerably pending on the expected result of the inventive method. In some aspects, it may be sufficient that the library comprises two, three, or four different complexes. However, in most events, more than two different complexes are desired to obtain a higher diversity. In some aspects, the library comprises 1,000 or more different complexes, more preferred 1,000,000 or more different complexes. The upper limit for the size of the library is only restricted by the size of the vessel in which the library is comprised. It may be calculated that a vial may comprise up to 10^{14} different complexes.

Methods for forming libraries of complexes

The complexes comprising an identifier having two or more codons that codes for reactants that have reacted in the formation of the molecule part of the complex may be formed by a variety of processes. Generally, the preferred methods can be used for the formation of virtually any kind of encode molecule. Suitable examples of processes include prior art methods disclosed in WO 93/20242, WO 93/06121, WO 00/23458, WO 02/074929, and WO 02/103008, the content of which being incorporated herein by reference as well as methods of the present applicant not yet public available, including the methods disclosed in DK PA 2002 01955 filed 19 December 2002, and DK PA 2003 00430 filed 20 March 2003. Any of these methods may be used, and the entire content of the patent applications are included herein by reference.

Below four preferred embodiments are described. A first embodiment disclosed in more detail in WO 02/103008 is based on the use of a polymerase to incorporate unnatural nucleotides as building blocks. Initially, a plurality of template oligonucleotides is provided. Subsequently primers are annealed to each of the templates and a polymerase is extending the primer using nucleotide derivatives which have appended chemical entities. Subsequent to or simultaneously with the incorporation of the nucleotide derivatives, the chemical entities are reacted to form a reaction product. The encoded molecule may be post-modified by cleaving some of the linking moieties to better present the encoded molecule.

Several possible reaction approaches for the chemical entities are apparent. First, the nucleotide derivatives can be incorporated and the chemical entities subsequently polymerised. In the event the chemical entities each carry two reactive groups, the chemical entities can be attached to adjacent chemical entities by a reaction of these reactive groups. Exemplary of the reactive groups are amine and carboxylic acid, which upon reaction form an amide bond. Adjacent chemical entities can also be linked together using a linking or bridging moiety. Exemplary of this approach is the linking of two chemical entities each bearing an amine group by a bi-carboxylic acid. Yet another approach is the use of a reactive group between a chemical entity and the nucleotide building block, such as an ester or a thioester group. An adjacent building block having a reactive group such as an amine may

cleave the interspaced reactive group to obtain a linkage to the chemical entity, e.g. by an amide linking group.

A second embodiment for obtainment of complexes pertains to the use of hybridisation of building blocks to a template and reaction of chemical entities attached to the building blocks in order to obtain a reaction product. This approach comprises that templates are contacted with a plurality of building blocks, wherein each building block comprises an anti-codon and a chemical entity. The anti-codons are designed such that they recognise a sequence, i.e. a codon, on the template. Subsequent to the annealing of the anti-codon and the codon to each other a reaction of the chemical entity is effected.

The template may be associated with a scaffold. Building blocks bringing chemical entities in may be added sequentially or simultaneously and a reaction of the reactive group of the chemical entity may be effected at any time after the annealing of the building blocks to the template.

A third embodiment for the generation of a complex includes chemical or enzymatical ligation of building blocks when these are lined up on a template. Initially, templates are provided, each having one or more codons. The templates are contacted with building blocks comprising anti-codons linked to chemical entities. The two or more anti-codons annealed on a template are subsequently ligated to each other and a reaction of the chemical entities is effected to obtain a reaction product. The method is disclosed in more detail in DK PA 2003 00430 filed 20 March 2003.

A fourth embodiment makes use of the extension by a polymerase of an affinity sequence of the nascent complex to transfer the anti-codon of a building block to the nascent complex. The method implies that a nascent complex comprising a scaffold and an affinity region is annealed to a building block comprising a region complementary to the affinity section. Subsequently the anti-codon region of the building block is transferred to the nascent complex by a polymerase. The transfer of the chemical entity may be transferred prior to, simultaneously with or subsequent to the transfer of the anti-codon. This method is disclosed in detail in DK PA 2002 01955 filed 19 December 2002.

Thus, the codons are either pre-made into one or more templates before the encoded molecules are generated or the codons are transferred simultaneously with the formation of the encoded molecules.

After or simultaneously with the formation of the reaction product some of the linkers to the template may be cleaved, however at least one linker must be maintained to provide for the complex.

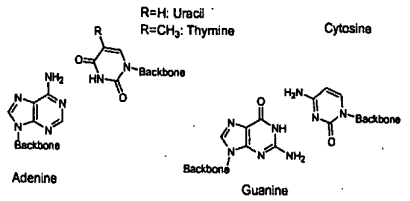
Nucleotides

The nucleotides used in the present invention may be linked together in a sequence of nucleotides, i.e. an oligonucleotide. Each nucleotide monomer is normally composed of two parts, namely a nucleobase moiety, and a backbone. The back bone may in some cases be subdivided into a sugar moiety and an internucleoside linker.

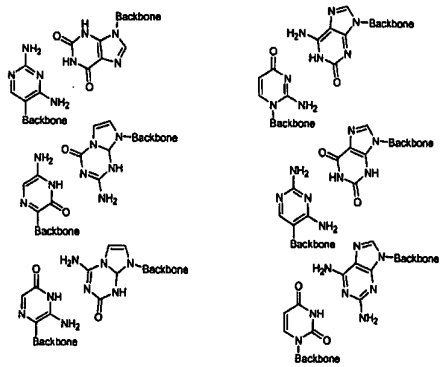
The nucleobase moiety may be selected among naturally occurring nucleobases as well as non-naturally occurring nucleobases. Thus, "nucleobase" includes not only the known purine and pyrimidine hetero-cycles, but also heterocyclic analogues and tautomers thereof. Illustrative examples of nucleobases are adenine, guanine, thymine, cytosine, uracil, purine, xanthine, diaminopurine, 8-oxo-N^6-methyladenine, 7-deazaxanthine, 7-deazaguanine, N^4,N^4-ethanocytosin, N^6,N^6-ethano-2,6-diamino-purine, 5-methylcytosine, 5-(C^3-C^6)-alkynylcytosine, 5-fluorouracil, 5-bromouracil, pseudoisocytosine, 2-hydroxy-5-methyl-4-triazolopyridine, isocytosine, isoguanine, inosine and the "non-naturally occurring" nucleobases described in Benner et al., U.S. Pat No. 5,432,272. The term "nucleobase" is intended to cover these examples as well as analogues and tautomers thereof. Especially interesting nucleobases are adenine, guanine, thymine, cytosine, 5-methylcytosine, and uracil, which are considered as the naturally occurring nucleobases in relation to therapeutic and diagnostic application in humans.

Examples of suitable specific pairs of nucleobases are shown below:

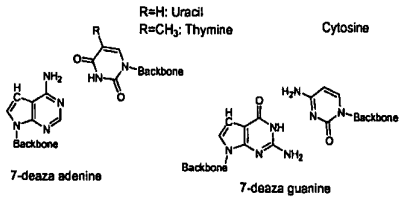
Natural Base Pairs



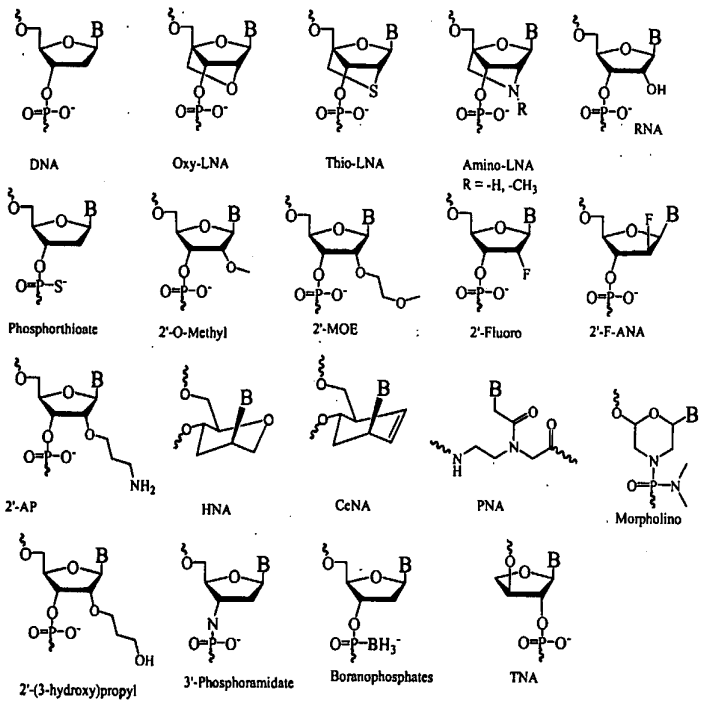
Synthetic Base Pairs



Synthetic purine bases pairing with natural pyrimidines



Suitable examples of backbone units are shown below (B denotes a nucleobase):

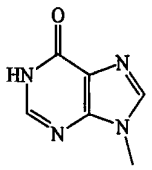


The sugar moiety of the backbone is suitably a pentose but may be the appropriate part of an PNA or a six-member ring. Suitable examples of possible pentoses include ribose, 2'-deoxyribose, 2'-O-methyl-ribose, 2'-flour-ribose, and 2'-4'-O-methylene-ribose (LNA). Suitably the nucleobase is attached to the 1' position of the pentose entity.

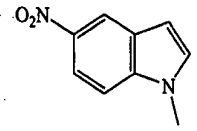
An internucleoside linker connects the 3' end of preceding monomer to a 5' end of a succeeding monomer when the sugar moiety of the backbone is a pentose, like ribose or 2-deoxyribose. The internucleoside linkage may be the natural occurring phospodiester linkage or a derivative thereof. Examples of such derivatives include phosphorothioate, methylphosphonate, phosphoramidate, phosphotriester, and phosphodithioate. Furthermore, the internucleoside linker can be any of a number of non-phosphorous-containing linkers known in the art.

Preferred nucleic acid monomers include naturally occurring nucleosides forming part of the DNA as well as the RNA family connected through phosphodiester linkages. The members of the DNA family include deoxyadenosine, deoxyguanosine, deoxythymidine, and deoxycytidine. The members of the RNA family include adenosine, guanosine, uridine, cytidine, and inosine. Inosine is a non-specific pairing nucleoside and may be used as universal base because inosine can pair nearly isoenergetically with A, T, and C. Other compounds having the same ability of non-specifically base-pairing with natural nucleobases have been formed. Suitable compounds which may be utilized in the present invention includes among others the compounds depicted below

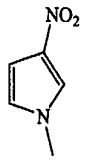
Examples of Universal Bases:



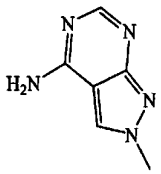
Inosine



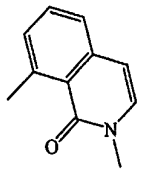
5-Nitroindole



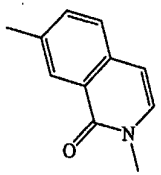
3-Nitropyrrole



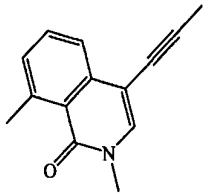
N^8-8aza-7deazaadenine



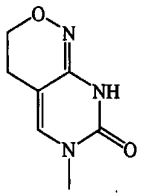
MICS



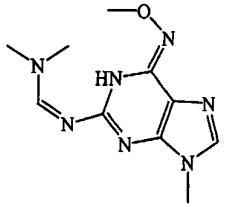
5MICS



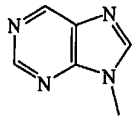
PIM



dP



dK



Nebularine

Building block

The chemical entities that are precursors for structural additions or eliminations of the encoded molecule may be attached to a building block prior to the participation in the formation of the reaction product leading the final encoded molecule. Besides the chemical entity, the building block generally comprises an anti-codon.

The chemical entity of the building block comprises at least one reactive group capable of participating in a reaction which results in a connection between the chemical entity of the building block and another chemical entity or a scaffold associated with the nascent complex. The connection is facilitated by one or more reactive groups of the chemical entity. The number of reactive groups which appear on the chemical entity is suitably one to ten. A building block featuring only one reactive group is used i.a. in the end positions of polymers or scaffolds, whereas building blocks having two reactive groups are suitable for the formation of the body part of a polymer or scaffolds capable of being reacted further. One, two or more reactive groups intended for the formation of connections, are typically present on scaffolds.

The reactive group of the building block may be capable of forming a direct connection to a reactive group of the nascent complex or the reactive group of the building block may be capable of forming a connection to a reactive group of the nascent complex through a bridging fill-in group. It is to be understood that not all the atoms of a reactive group are necessarily maintained in the connection formed. Rather, the reactive groups are to be regarded as precursors for the structure of the connection.

The subsequent cleavage step to release the chemical entity from the building block can be performed in any appropriate way. In an aspect of the invention the cleavage involves usage of a reagent or an enzyme. The cleavage results in a transfer of the chemical entity to the nascent encoded molecule or in a transfer of the nascent encoded molecule to the chemical entity of the building block. In some cases it may be advantageous to introduce new chemical groups as a consequence of linker cleavage. The new chemical groups may be used for further reaction in a subsequent cycle, either directly or after having been activated. In other cases it is desirable that no trace of the linker remains after the cleavage.

In another aspect, the connection and the cleavage is conducted as a simultaneous reaction, i.e. either the chemical entity of the building block or the nascent encoded molecule is a leaving group of the reaction. In general, it is preferred to design the system such that the connection and the cleavage occur simultaneously because this will reduce the number of steps and the complexity. The simultaneous connection and cleavage can also be designed such that either no trace of the linker remains or such that a new chemical group for further reaction is introduced, as described above.

The attachment of the chemical entity to the building block, optionally via a suitable spacer can be at any entity available for attachment, e.g. the chemical entity can be attached to a nucleobase or the backbone. In general, it is preferred to attach the chemical entity at the phosphor of the internucleoside linkage or at the nucleobase. When the nucleobase is used for attachment of the chemical entity, the attachment point is usually at the 7 position of the purines or 7-deaza-purins or at the 5 position of pyrimidines. The nucleotide may be distanced from the reactive group of the chemical entity by a spacer moiety. The spacer may be designed such that the conformational space sampled by the reactive group is optimized for a reaction with the reactive group of the nascent encoded molecule or reactive site.

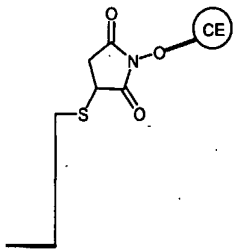
The anticodon complements the codon of the identifier sequence and generally comprises the same number of nucleotides as the codon. The anticodon may be adjoined with a fixed sequence, such as a sequence complementing a framing sequence.

Various specific building blocks are envisaged. Building blocks of particular interest are shown below.

<u>Building blocks transferring a chemical entity to a recipient nucleophilic group</u>

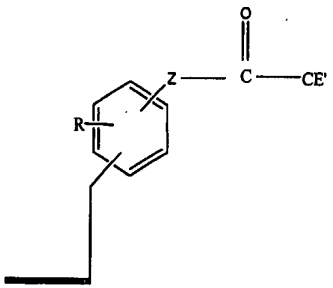
The building block indicated below is capable of transferring a chemical entity (CE) to a recipient nucleophilic group, typically an amine group. The bold lower horizontal line illustrates the building block and the vertical line illustrates a spacer. The 5-membered substituted N-hydroxysuccinimid (NHS) ring serves as an activator, i.e. a labile bond is formed between the oxygen atom connected to the NHS ring and the

chemical entity. The labile bond may be cleaved by a nucleophilic group, e.g. positioned on a scaffold



The 5-membered substituted N-hydroxysuccinimid (NHS) ring serves as an activator, i.e. a labile bond is formed between the oxygen atom connected to the NHS ring and the chemical entity. The labile bond may be cleaved by a nucleophilic group, e.g. positioned on a scaffold, to transfer the chemical entity to the scaffold, thus converting the remainder of the fragment into a leaving group of the reaction. When the chemical entity is connected to the activator through an carbonyl group and the recipient group is an amine, the bond formed on the scaffold will an amide bond. The above building block is the subject of the Danish patent application No. PA 2002 01946 and the US provisional patent application No. 60/434,439, the content of which are incorporated herein in their entirety by reference.

Another building block which may form an amide bond is



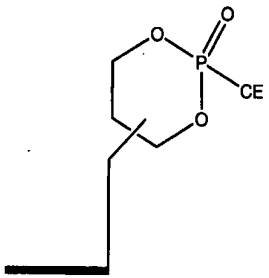
R may be absent or NO_2, CF_3, halogen, preferably Cl, Br, or I, and Z may be S or O. This type of building block is disclosed in Danish patent application No. PA 2002 0951 and US provisional patent application filed 20 December 2002 with the title "A

building block capable of transferring a functional entity to a recipient reactive group". The content of both patent application are incorporated herein in their entirety by reference.

A nucleophilic group can cleave the linkage between Z and the carbonyl group thereby transferring the chemical entity –(C=O)-CE' to said nucleophilic group.

Building blocks transferring a chemical entity to a recipient reactive group forming a C=C bond

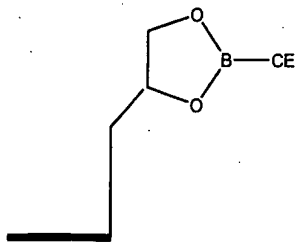
A building block as shown below are able to transfer the chemical entity to a recipient aldehylde group thereby forming a double bond between the carbon of the aldehyde and the chemical entity



The above building block is comprised by the Danish patent application No. DK PA 2002 01952 and the US provisional patent application filed 20 December 2002 with the title "A building block capable of transferring a functional entity to a recipient reactive group forming a C=C double bond". The content of both patent applications are incorporated herein in their entirety by reference.

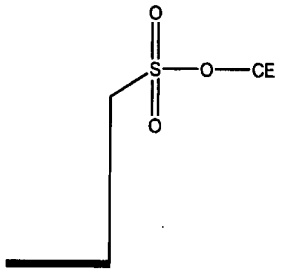
Building blocks transferring a chemical entity to a recipient reactive group forming a C-C bond

The below building block is able to transfer the chemical entity to a recipient group thereby forming a single bond between the receiving moiety, e.g. a scaffold, and the chemical entity.



The above building block is comprised by the Danish patent application No. DK PA 2002 01947 and the US provisional patent application No 60/434,428. The content of both patent applications are incorporated herein in their entirety by reference.

Another building block capable of transferring a chemical entity to a receiving reactive group forming a single bond is



The receiving group may be a nucleophile, such as a group comprising a hetero atom, thereby forming a single bond between the chemical entity and the hetero atom, or the receiving group may be an electronegative carbon atom, thereby forming a C-C bond between the chemical entity and the scaffold.

The chemical entity attached to any of the above building blocks may be a selected from a large arsenal of chemical structures. Examples of chemical entities are H or entities selected among the group consisting of a C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_4-C_8 alkadienyl, C_3-C_7 cycloalkyl, C_3-C_7 cycloheteroalkyl, aryl, and heteroaryl, said group being substituted with 0-3 R^4, 0-3 R^5 and 0-3 R^9 or C_1-C_3 alkylene-NR^4_2, C_1-C_3 alkylene-$NR^4C(O)R^8$, C_1-C_3 alkylene-$NR^4C(O)OR^8$, C_1-C_2 al-

kylene-O-NR^4_2, C_1-C_2 alkylene-O-$NR^4C(O)R^8$, C_1-C_2 alkylene-O-$NR^4C(O)OR^8$ substituted with 0-3 R^9.

where R^4 is H or selected independently among the group consisting of C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_3-C_7 cycloalkyl, C_3-C_7 cycloheteroalkyl, aryl, heteroaryl, said group being substituted with 0-3 R^9 and

R^5 is selected independently from $-N_3$, -CNO, $-C(NOH)NH_2$, -NHOH, $-NHNHR^6$, $-C(O)R^8$, $-SnR^6_3$, $-B(OR^6)_2$, $-P(O)(OR^6)_2$ or the group consisting of C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_4-C_8 alkadienyl said group being substituted with 0-2 R^7,

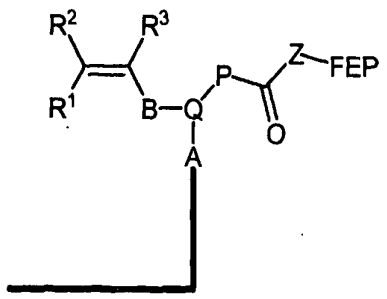
where R^6 is selected independently from H, C_1-C_6 alkyl, C_3-C_7 cycloalkyl, aryl or C_1-C_6 alkylene-aryl substituted with 0-5 halogen atoms selected from -F, -Cl, -Br, and –I; and R^7 is independently selected from $-NO_2$, $-COOR^6$, $-COR^6$, -CN, $-OSiR^6_3$, $-OR^6$ and $-NR^6_2$.

R^8 is H, C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_3-C_7 cycloalkyl, aryl or C_1-C_6 alkylene-aryl substituted with 0-3 substituents independently selected from –F, -Cl, $-NO_2$, $-R^3$, $-OR^3$, $-SiR^3_3$

R^9 is =O, -F, -Cl, -Br, -I, -CN, $-NO_2$, $-OR^6$, $-NR^6_2$, $-NR^6$-$C(O)R^8$, $-NR^6$-$C(O)OR^8$, $-SR^6$, $-S(O)R^6$, $-S(O)_2R^6$, $-COOR^6$, $-C(O)NR^6_2$ and $-S(O)_2NR^6_2$.

Cross-link cleavage building blocks

It may be advantageous to split the transfer of a chemical entity to a recipient reactive group into two separate steps, namely a cross-linking step and a cleavage step because each step can be optimized. A suitable building block for this two step process is illustrated below:



Initally, a reactive group appearing on the functional entity precursor (abbreviated FEP) reacts with a recipient reactive group, e.g. a reactive group appearing on a

scaffold, thereby forming a cross-link. Subsequently, a cleavage is performed, usually by adding an aqueous oxidising agent such as I_2, Br_2, Cl_2, H^+, or a Lewis acid. The cleavage results in a transfer of the group HZ-FEP- to the recipient moiety, such as a scaffold.

In the above formula

Z is O, S, NR^4

Q is N, CR^1

P is a valence bond, O, S, NR^4, or a group C_{5-7}arylene, C_{1-6}alkylene, C_{1-6}O-alkylene, C_{1-6}S-alkylene, NR^1-alkylene, C_{1-6}alkylene-O, C_{1-6}alkylene-S option said group being substituted with 0-3 R^4, 0-3 R^5 and 0-3 R^9 or C_1-C_3 alkylene-NR^4_2, C_1-C_3 alkylene-$NR^4C(O)R^8$, C_1-C_3 alkylene-$NR^4C(O)OR^8$, C_1-C_2 alkylene-O-NR^4_2, C_1-C_2 alkylene-O-$NR^4C(O)R^8$, C_1-C_2 alkylene-O-$NR^4C(O)OR^8$ substituted with 0-3 R^9,

B is a group comprising D-E-F, in which

D is a valence bond or a group C_{1-6}alkylene, C_{1-6}alkenylene, C_{1-6}alkynylene, C_{5-7}arylene, or C_{5-7}heteroarylene, said group optionally being substituted with 1 to 4 group R^{11},

E is, when present, a valence bond, O, S, NR^4, or a group C_{1-6}alkylene, C_{1-6}alkenylene, C_{1-6}alkynylene, C_{5-7}arylene, or C_{5-7}heteroarylene, said group optionally being substituted with 1 to 4 group R^{11},

F is, when present, a valence bond, O, S, or NR^4,

A is a spacing group distancing the chemical structure from the complementing element, which may be a nucleic acid,

R^1, R^2, and R^3 are independent of each other selected among the group consisting of H, C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_4-C_8 alkadienyl, C_3-C_7 cycloalkyl, C_3-C_7 cycloheteroalkyl, aryl, and heteroaryl, said group being substituted with 0-3 R^4, 0-3 R^5 and 0-3 R^9 or C_1-C_3 alkylene-NR^4_2, C_1-C_3 alkylene-$NR^4C(O)R^8$, C_1-C_3 alkylene-$NR^4C(O)OR^8$, C_1-C_2 alkylene-O-NR^4_2, C_1-C_2 alkylene-O-$NR^4C(O)R^8$, C_1-C_2 alkylene-O-$NR^4C(O)OR^8$ substituted with 0-3 R^9,

FEP is a group selected among the group consisting of H, C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_4-C_8 alkadienyl, C_3-C_7 cycloalkyl, C_3-C_7 cycloheteroalkyl, aryl, and heteroaryl, said group being substituted with 0-3 R^4, 0-3 R^5 and 0-3 R^9 or C_1-C_3 alkylene-NR^4_2, C_1-C_3 alkylene-$NR^4C(O)R^8$, C_1-C_3 alkylene-$NR^4C(O)OR^8$,

C_1-C_2 alkylene-O-NR^4_2, C_1-C_2 alkylene-O-$NR^4C(O)R^8$, C_1-C_2 alkylene-O-$NR^4C(O)OR^8$ substituted with 0-3 R^9,

where R^4 is H or selected independently among the group consisting of C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_3-C_7 cycloalkyl, C_3-C_7 cycloheteroalkyl, aryl, heteroaryl, said group being substituted with 0-3 R^9 and

R^5 is selected independently from $-N_3$, -CNO, $-C(NOH)NH_2$, -NHOH, $-NHNHR^6$, $-C(O)R^6$, $-SnR^6_3$, $-B(OR^6)_2$, $-P(O)(OR^6)_2$ or the group consisting of C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_4-C_8 alkadienyl said group being substituted with 0-2 R^7,

where R^6 is selected independently from H, C_1-C_6 alkyl, C_3-C_7 cycloalkyl, aryl or C_1-C_6 alkylene-aryl substituted with 0-5 halogen atoms selected from -F, -Cl, -Br, and –I; and R^7 is independently selected from $-NO_2$, $-COOR^6$, $-COR^6$, -CN, $-OSiR^6_3$, $-OR^6$ and $-NR^6_2$.

R^8 is H, C_1-C_6 alkyl, C_2-C_6 alkenyl, C_2-C_6 alkynyl, C_3-C_7 cycloalkyl, aryl or C_1-C_6 alkylene-aryl substituted with 0-3 substituents independently selected from –F, -Cl, $-NO_2$, $-R^3$, $-OR^3$, $-SiR^3_3$

R^9 is =O, -F, -Cl, -Br, -I, -CN, $-NO_2$, $-OR^6$, $-NR^6_2$, $-NR^6$-$C(O)R^8$, $-NR^6$-$C(O)OR^8$, $-SR^6$, $-S(O)R^6$, $-S(O)_2R^6$, $-COOR^6$, $-C(O)NR^6_2$ and $-S(O)_2NR^6_2$.

In a preferred embodiment Z is O or S, P is a valence bond, Q is CH, B is CH_2, and R^1, R^2, and R^3 is H. The bond between the carbonyl group and Z is cleavable with aqueous I_2.

Contacting between target and library

The contacting step, by which the library of bifunctional molecules is subjected under binding conditions to a target, may be referred to as the enrichment step or the selection step, as appropriate, and includes the screening of the library for synthetic molecules having predetermined desirable characteristics. Predetermined desirable characteristics can include binding to a target, catalytically changing the target, chemically reacting with a target in a manner which alters/modifies the target or the functional activity of the target, and covalently attaching to the target as in a suicide inhibitor.

In theory, molecules of interest can be selected based on their properties using either physical or physiological procedures. The method preferred according to the present invention is to enrich molecules with respect to binding affinity towards a

target of interest. In a certain embodiment, the basic steps involve mixing the library of complexes with the immobilized target of interest. The target can be attached to a column matrix or microtitre wells with direct immobilization or by means of antibody binding or other high-affinity interactions. In another embodiment, the target and displayed molecules interact without immobilisation of the target. Displayed molecules that bind to the target will be retained on this surface, while nonbinding displayed molecules will be removed during a single or a series of wash steps. The identifiers of complexes bound to the target can then be separated by cleaving the physical connection to the synthetic molecule. It may be considered advantageously to perform a chromatography step after of instead of the washing step. After the cleavage of the physical link between the synthetic molecule and the identifier, the identifier may be recovered from the media and optionally amplified before the decoding step.

A significant reduction in background binders may be obtained with increased washing volumes, repeating washing steps, higher detergent concentrations and prolonged incubation during washing. Thus, the more volume and number of steps used in the washing procedure together with more stringent conditions will more efficiently remove non-binders and background binders. The right stringency in the washing step can also be used to remove low-affinity specific binders. However, the washing step will also remove wanted binders if too harsh conditions are used.

A blocking step, such as incubation of solid phase with skimmed milk proteins or other inert proteins and/or mild detergent such as Tween-20 and Triton X-100, may also be used to reduce the background. The washing conditions should be as stringent as possible to remove background binding but to retain specific binders that interact with the immobilized target. Generally, washing conditions are adjusted to maintain the desired affinity binders, e.g. binders in the micromolar, nanomolar, or pocomolar range.

In traditional elution protocols, false positives due to suboptimal binding and washing conditions are difficult to circumvent and may require elaborate adjustments of experimental conditions. However, an enrichment of more than 100 to 1000 is rarely obtained. The present invention alleviates the problem with false positive being obtained because the non-specific binding complexes to a large extent remain in the

reaction chamber. The experiments reported herein suggest that an enrichment of more than 10^7 can be obtained.

The target can be any compound of interest. E.g. the target can be a protein, peptide, carbohydrate, polysaccharide, glycoprotein, hormone, receptor, antigen, antibody, virus, substrate, metabolite, transition state analogue, cofactor, inhibitor, drug, dye, nutrient, growth factor, cell, tissue, etc. without limitation. Suitable targets include, but are not limited to, angiotensin converting enzyme, renin, cyclooxygenase, 5-lipoxygenase, IIL- 1 0 converting enzyme, cytokine receptors, PDGF receptor, type II inosine monophosphate dehydrogenase, β-lactamases, integrin, proteases like factor VIIa, kinases like Bcr-Abl/Her, phosphotases like PTP-1B, and fungal cytochrome P-450. Targets can include, but are not limited to, bradykinin, neutrophil elastase, the HIV proteins, including *tat, rev, gag, int,* RT, nucleocapsid etc., VEGF, bFGF, TGFβ, KGF, PDGF, GPCR, thrombin, substance P, IgE, sPLA2, red blood cells, glioblastomas, fibrin clots, PBMCs, hCG, lectins, selectins, cytokines, ICP4, complement proteins, etc.

A target can also be a surface of a non-biological origin, such as a polymer surface or a metal surface. The method of the invention may then be used to identify suitable coatings for such surfaces.

In a preferred embodiment, the desirable synthetic molecule acts on the target without any interaction between the nucleic acid attached to the desirable encoded molecule and the target. In one embodiment, the bound complex-target aggregate can be partitioned from unbound complexes by a number of methods. The methods include nitrocellulose filter binding, column chromatography, filtration, affinity chromatography, centrifugation, and other well known methods. A preferred method is size-exclusion chromatography.

Briefly, the library of complexes is subjected to the target, which may include contact between the library and a column onto which the target is immobilised. Identifiers associated with undesirable synthetic molecules, i.e. synthetic molecules not bound to the target under the stringency conditions used, will pass through the column. Additional undesirable synthetic molecules (e.g. synthetic molecules which cross-react with other targets) may be removed by counter-selection methods. Desirable

complexes are bound to the column. The target may be immobilized in a number of ways. In one embodiment, the target is immobilized through a cleavable physical link, such as one more chemical bonds. The aggregate of the target and the complex may then be subjected to a size exclusion chromatography to separate the aggregate from the rest of the compounds in the media. The complex may then be eluted from the target by changing the conditions (e.g., salt, pH, surfactant, temperature etc.). Alternatively, the complex may be provided with a cleavable linker, preferable orthogonal to the cleavable linker that attached the target to the solid support, at a position between the synthetic molecule and the identifier. Subsequent to the size exclusion chromatography this cleavable linker is cleaved to separate the identifiers of complexes having affinity towards the targets. Just to mention a single type of orthogonal cleavable linkages, one could attached to target to the solid support through a linkage that can be cleaved by a chemical agent, and the linker separating the synthetic molecule and the identifier may be selected as a photocleavable linkage. More specifically, the former linkage may be a disulphide bond that can be cleaved by a suitable reducing agent like DTT (dithiothreitol) and the latter linkage may be a *o*-nitrophenyl group.

There are other partitioning and screening processes which are compatible with this invention that are known to one of ordinary skill in the art. In one embodiment, the products can be fractionated by a number of common methods and then each fraction is then assayed for activity. The fractionization methods can include size, pH, hydrophobicity, etc.

Inherent in the present method is the selection of encoded molecules on the basis of a desired function; this can be extended to the selection of molecules with a desired function and specificity. Specificity can be required during the selection process by first extracting complexes which are capable of interacting with a non-desired "target" (negative selection, or counter-selection), followed by positive selection with the desired target. As an example, inhibitors of fungal cytochrome P-450 are known to cross-react to some extent with mammalian cytochrome P-450 (resulting in serious side effects). Highly specific inhibitors of the fungal cytochrome could be selected from a library by first removing those complexes capable of interacting with the mammalian cytochrome, followed by retention of the remaining products which are

capable of interacting with the fungal cytochrome.

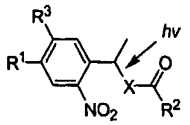
Cleavable linkers

A cleavable linker may be positioned between the target and a solid support or between the potential drug candidate and the identifier or any other position that may ensure a separation of the identifier from successful complexes from non-specific binding complexes. The cleavable linker may be selectively cleavable, i.e. conditions may selected that only cleave that particular linker.

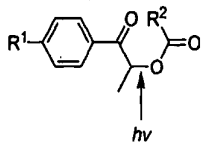
The cleavable linkers may be selected from a large plethora of chemical structures. Examples of linkers includes, but are not limited to, linkers having an enzymatic cleavage site, linkers comprising a chemical degradable component, linkers cleavable by electromagnetic radiation.

Examples of linkers cleavable by electromagnetic radiation (light)

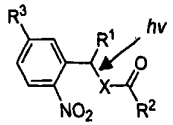
o-nitrobenzyl



p-alkoxy

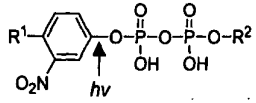


o-nitrobenzyl in exo position



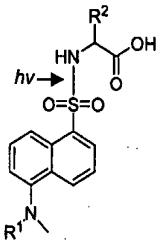
For more details see Holmes CP. J. Org. Chem. **1997,** *62*, 2370-2380

3-nitrophenyloxy



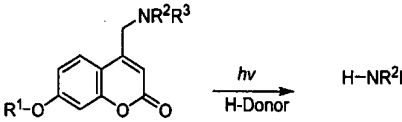
For more details see Rajasekharan Pillai, V. N. Synthesis. **1980**, 1-26

Dansyl derivatives:



For more details see Rajasekharan Pillai, V. N. Synthesis. **1980**, 1-26

Coumarin derivatives



For more details see R. O. Schoenleber, B. Giese. Synlett **2003**, 501-504

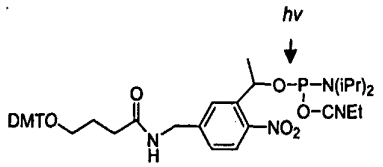
R^1 and R^2 can be either of the potential drug candidate and the identifier, respectively. Alternatively, R^1 and R^2 can be either of the target or a solid support, respectively.

R^3 = H or OCH_3

If X is O then the product will be a carboxylic acid

If X is NH the product will be a carboxamide

One specific example is the PC Spacer Phosphoramidite (Glen research catalog # 10-4913-90) which can be introduced in an oligonucleotide during synthesis and cleaved by subjecting the sample in water to UV light (~ 300-350 nm) for 30 seconds to 1 minute.

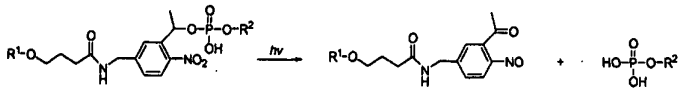


DMT = 4,4'-Dimethoxytrityl

iPr = Isopropyl

CNEt = Cyanoethyl

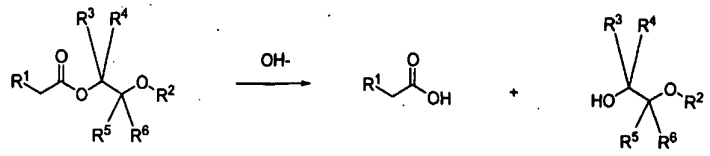
The above PC spacer phosphoamidite is suitable incorporated in a library of complexes at a position between the indentifier and the potential drug candidate. The spacer may be cleaved according to the following reaction.



R^1 and R^2 can be either of the encoded molecule and the identifying molecule, respectively. In a preferred aspect R^2 is an oligonucleotide identifier and the R^1 is the potential drug candidate. When the linker is cleaved a phosphate group is generated allowing for further biological reactions. As an example, the phosphate group may be positioned in the 5´end of an oligonucleotide allowing for an enzymatic ligation process to take place.

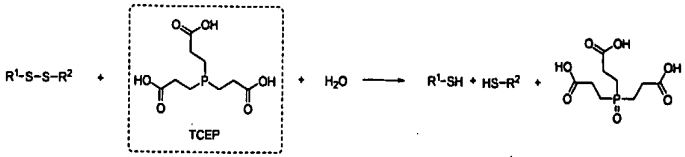
Examples of linkers cleavable by chemical agents:

Ester linkers can be cleaved by nucleophilic attack using e.g. hydroxide ions. In practice this can be accomplished by subjecting the target-ligand complex to a base for a short period.



R^1 and R^2 can be the either of be the potential drug candidate or the identifier, respectively. R^{4-8} can be any of the following: H, CN, F, NO_2, SO_2NR_2.

Disulfide linkers can efficiently be cleaved / reduced by Tris (2-carboxyethyl) phosphine (TCEP). TCEP selectively and completely reduces even the most stable water-soluble alkyl disulfides over a wide pH range. These reductions frequently required less than 5 minutes at room temperature. TCEP is a non-volatile and odorless reductant and unlike most other reducing agents, it is resistant to air oxidation. Trialkylphosphines such as TCEP are stable in aqueous solution, selectively reduce disulfide bonds, and are essentially unreactive toward other functional groups commonly found in proteins.



More details on the reduction of disulfide bonds can be found in Kirley, T.L.(1989), Reduction and fluorescent labeling of cyst(e)ine-containing proteins for subsequent structural analysis, *Anal. Biochem.* **180**, 231 and Levison, M.E., *et al.* (1969), Reduction of biological substances by water-soluble phosphines: Gamma-globulin. *Experentia* **25**, 126-127.

Linkers cleavable by enzymes

The linker connecting the potential drug candidate with the identifier or the solid support and the target can include a peptide region that allows a specific cleavage using a protease. This is a well-known strategy in molecular biology. Site-specific proteases and their cognate target amino acid sequences are often used to remove

the fusion protein tags that facilitate enhanced expression, solubility, secretion or purification of the fusion protein.

Various proteases can be used to accomplish a specific cleavage. The specificity is especially important when the cleavage site is presented together with other sequences such as for example the fusion proteins. Various conditions have been optimized in order to enhance the cleavage efficiency and control the specificity. These conditions are available and know in the art.

Enterokinase is one example of an enzyme (serine protease) that cut a specific amino acid sequence. Enterokinase recognition site is Asp-Asp-Asp-Asp-Lys (DDDDK), and it cleaves C-terminally of Lys. Purified recombinant Enterokinase is commercially available and is highly active over wide ranges in pH (pH 4.5-9.5) and temperature (4-45°C).

The nuclear inclusion protease from tobacco etch virus (TEV) is another commercially available and well-characterized proteases that can be used to cut at a specific amino acid sequence. TEV protease cleaves the sequence Glu-Asn-Leu-Tyr-Phe-Gln-Gly/Ser (ENLYFQG/S) between Gln-Gly or Gln-Ser with high specificity.

Another well-known protease is thrombin that specifically cleaves the sequence Leu-Val-Pro-Arg-Gly-Ser (LVPAGS) between Arg-Gly. Thrombin has also been used for cleavage of recombinant fusion proteins. Other sequences can also be used for thrombin cleavage; these sequences are more or less specific and more or less efficiently cleaved by thrombin. Thrombin is a highly active protease and various reaction conditions are known to the public.

Activated coagulation factor FX (FXa) is also known to be a specific and useful protease. This enzyme cleaves C-terminal of Arg at the sequence Ile-Glu-Gly-Arg (IEGR). FXa is frequently used to cut between fusion proteins when producing proteins with recombinant technology. Other recognition sequences can also be used for FXa.

Other types of proteolytic enzymes can also be used that recognize specific amino acid sequences. In addition, proteolytic enzymes that cleave amino acid sequences

in an un-specific manner can also be used if only the linker contains an amino acid sequence in the complex molecule.

Other type of molecules such as ribozymes, catalytically active antibodies, or lipases can also be used. The only prerequisite is that the catalytically active molecule can cleave the specific structure used as the linker, or as a part of the linker, that connects the encoding region and the displayed molecule or, in the alternative the solid support and the target.

A variety of endonucleases are available that recognize and cleave a double stranded nucleic acid having a specific sequence of nucleotides. The endonuclease Eco RI is an example of a nuclease that efficiently cuts a nucleotide sequence linker comprising the sequence GAATTC also when this sequence is close to the nucleotide sequence length. Purified recombinant Eco RI is commercially available and is highly active in a range of buffer conditions. As an example the Eco RI is working in in various protocols as indicted below (NEBuffer is available from New England Biolabs):

NEBuffer 1 : [10 mM Bis Tris Propane-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol (pH 7.0 at 25°C)],

NEBuffer 2 : [50 mM NaCl, 10 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol (pH 7.9 at 25°C)],

NEBuffer 3 : [100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, 1 mM dithiothreitol (pH 7.9 at 25°C)],

NEBuffer 4 : [50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 1 mM dithiothreitol (pH 7.9 at 25°C)].

Extension buffer : mM KCl, 20 mM Tris-HCl(Ph 8.8 at 25° C), 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$ and 0.1% Triton X-100, and 200 µM dNTPs.

Determining the identifier sequence

The nucleotide sequence of the identifier sequence present in the isolated bifunctional molecules or the separated identifiers is determined to identify the chemical entities that participated in the binding interaction. The synthesis method of the synthetic molecule may be established if information on the chemical entities as well as the point in time they have been incorporated in the synthetic molecule can be

deduced from the identifier. It may be sufficient to get information on the chemical structure of the various chemical entities that have participated in the synthetic molecule to deduce the full molecule due to structural constraints during the formation. As an example, the use of different kinds of attachment chemistries may ensure that a chemical entity on a building block can only be transferred to a single position on a scaffold. Another kind of chemical constrains may be present due to steric hindrance on the scaffold molecule or the chemical entity to be transferred. In general however, it is preferred that information can be inferred from the identifier sequence that enable the identification of each of the chemical entities that have participated in the formation of the encoded molecule along with the point in time in the synthesis history the chemical entities have been incorporated in the (nascent) synthetic molecule.

Although conventional DNA sequencing methods are readily available and useful for this determination, the amount and quality of isolated bifunctional molecule may require additional manipulations prior to a sequencing reaction.

Where the amount is low, it is preferred to increase the amount of the identifier sequence by polymerase chain reaction (PCR) using PCR primers directed primer binding sites present in the identifier sequence.

In addition, the quality of the isolated bifunctional molecule may be such that multiple species of bifunctional molecule are co-isolated by virtue of similar capacities for binding to the target. In cases where more than one species of bifunctional molecule are isolated, the different isolated species must be separated prior to sequencing of the identifier oligonucleotide.

Thus in one embodiment, the different identifier sequences of the isolated bifunctional complexes are cloned into separate sequencing vectors prior to determining their sequence by DNA sequencing methods. This is typically accomplished by amplifying all of the different identifier sequences by PCR as described herein, and then using a unique restriction endonuclease sites on the amplified product to directionally clone the amplified fragments into sequencing vectors. The cloning and sequencing of the amplified fragments then is a routine procedure that can be carried out by any of a number of molecular biological methods known in the art.

Alternatively, the bifunctional complex or the PCR amplified identifier sequence can be analysed in a microarray. The array may be designed to analyse the presence of a single codon or multiple codons in a identifier sequence.

Synthesis of nucleic acids

Oligonucleotides can be synthesized by a variety of chemistries as is well known. For synthesis of an oligonucleotide on a substrate in the direction of 3' to 5', a free hydroxy terminus is required that can be conveniently blocked and deblocked as needed. A preferred hydroxy terminus blocking group is a dimexothytrityl ether (DMT). DMT blocked termini are first deblocked, such as by treatment with 3% dichloroacetic acid in dichloromethane (DCM) as is well known for oligonucleotide synthesis, to form a free hydroxy terminus.

Nucleotides in precursor form for addition to a free hydroxy terminus in the direction of 3' to 5' require a phosphoramidate moiety having an aminodiisopropyl side chain at the 3' terminus of a nucleotide. In addition, the free hydroxy of the phosphoramidate is blocked with a cyanoethyl ester (OCNET), and the 5' terminus is blocked with a DMT ether. The addition of a 5' DMT-, 3' OCNET-blocked phosphoramidate nucleotide to a free hydroxyl requires tetrazole in acetonitrile followed by iodine oxidation and capping of unreacted hydroxyls with acetic anhydride, as is well known for oligonucleotide synthesis. The resulting product contains an added nucleotide residue with a DMT blocked 5' terminus, ready for deblocking and addition of a subsequent blocked nucleotide as before.

For synthesis of an oligonucleotide in the direction of 5' to 3', a free hydroxy terminus on the linker is required as before. However, the blocked nucleotide to be added has the blocking chemistries reversed on its 5' and 3' termini to facilitate addition in the opposite orientation. A nucleotide with a free 3' hydroxyl and 5' DMT ether is first blocked at the 3' hydroxy terminus by reaction with TBS-Cl in imidazole to form a TBS ester at the 3' terminus. Then the DMT-blocked 5' terminus is deblocked with DCA in DCM as before to form a free 5' hydroxy terminus. The reagent (N,N-diisopropylamino)(cyanoethyl) phosphonamidic chloride having an aminodiisopropyl group and an OCNET ester is reacted in tetrahydrofuran (THF) with the 5' deblocked nucleotide to form the aminodiisopropyl-, OCNET-blocked phosphonami-

date group on the 5' terminus. Thereafter the 3' TBS ester is removed with tetrabutylammonium fluoride (TBAF) in DCM to form a nucleotide with the phosphonamidate-blocked 5' terminus and a free 3' hydroxy terminus. Reaction in base with DMT-Cl adds a DMT ether blocking group to the 3' hydroxy terminus.

The addition of the 3' DMT-, 5' OCNET-blocked phosphonamidated nucleotide to a linker substrate having a free hydroxy terminus then proceeds using the previous tetrazole reaction, as is well known for oligonucleotide polymerization. The resulting product contains an added nucleotide residue with a DMT-blocked 3' terminus, ready for deblocking with DCA in DCM and the addition of a subsequent blocked nucleotide as before.

Extension and amplification

The use of the polymerase chain reaction (PCR) is a preferred embodiment, for the production of the identifiers using the nucleic acids of the selected complexes as identifiers.

For use in this invention, the identifier sequences are preferably comprised of polynucleotide coding strands, such as mRNA and/or the sense strand of genomic DNA or non-natural nucleic acids, like TNA and LNA which may be used as template for a polymerase. If the genetic material to be processed is in the form of double stranded nucleic acid, it is usually first denatured, typically by melting, into single strands. The nucleic acid is subjected to a PCR reaction by treating (contacting) the sample with a PCR primer pair, each member of the pair having a preselected nucleotide sequence. The PCR primer pair is capable of initiating primer extension reactions by hybridizing to the PCR primer binding site on identifier oligonucleotide, preferably at least about 10 nucleotides in length, more preferably at least about 12 nucleotides in length. The first primer of a PCR primer pair is sometimes referred to as the "anti-sense primer" because it is extended into a non-coding or anti-sense strand of a nucleic acid, i.e., a strand complementary to a coding strand. The second primer of a PCR primer pair is sometimes referred to as the "sense primer" because it is adjoined with the coding or sense strand of a nucleic acid.

The PCR reaction is performed by mixing the PCR primer pair, preferably a predetermined amount thereof, with the nucleic acids of the sample, preferably a predetermined amount thereof, in a PCR buffer to form a PCR reaction admixture. The admixture is thermocycled for a number of cycles, which is typically predetermined, sufficient for the formation of a PCR reaction product, thereby amplifiyng the identifiers in the isolated complex.

PCR is typically carried out by thermocycling i.e., repeatedly increasing and decreasing the temperature of a PCR reaction admixture within a temperature range whose lower limit is about 30 degrees Celsius (30° C.) to about 55° C. and whose upper limit is about 90° C. to about 100° C. The increasing and decreasing can be continuous, but is preferably phasic with time periods of relative temperature stability at each of temperatures favoring polynucleotide synthesis, denaturation and hybridization.

A plurality of first primer and/or a plurality of second primers can be used in each amplification, e.g., one species of first primer can be paired with a number of different second primers to form several different primer pairs. Alternatively, an individual pair of first and second primers can be used. In any case, the amplification products of amplifications using the same or different combinations of first and second primers can be combined for assaying for mutations.

The PCR reaction is performed using any suitable method. Generally it occurs in a buffered aqueous solution, i.e., a PCR buffer, preferably at a pH of 7-9, most preferably about 8. Preferably, a molar excess of the primer is admixed to the buffer containing the identifier strand. A large molar excess is preferred to improve the efficiency of the process.

The PCR buffer also contains the deoxyribonucleotide triphosphates (polynucleotide synthesis substrates) dATP, dCTP, dGTP, and dTTP and a polymerase, typically thermostable, all in adequate amounts for primer extension (polynucleotide synthesis) reaction. The resulting solution (PCR admixture) is heated to about 90° C.-100° C. for about 1 to 10 minutes, preferably from 1 to 4 minutes. After this heating period the solution is allowed to cool to a primer hybridization temperature. The synthesis reaction may occur at from room temperature up to a temperature above which the polymerase (inducing agent) no longer functions efficiently. Thus, for example, if DNA polymerase is used as inducing agent, the temperature is generally no greater

than about 40° C. The thermocycling is repeated until the desired amount of PCR product is produced. An exemplary PCR buffer comprises the following: 50 mM KCl; 10 mM Tris-HCl at pH 8.3; 1.5 mM MgCl2 ; 0.001% (wt/vol) gelatin, 200 µM dATP; 200 µM dTTP; 200 µM dCTP; 200 µM dGTP; and 2.5 units Thermus aquaticus (Taq) DNA polymerase I (U.S. Pat. No. 4,889,818) per 100 microliters (µl) of buffer.

The inducing agent may be any compound or system which will function to accomplish the synthesis of primer extension products, including enzymes. Suitable enzymes for this purpose include, for example, E. coli DNA polymerase I, Klenow fragment of E. coli DNA polymerase I, T4 DNA polymerase, Taq DNA polymerase, Pfu polymerase, Vent polymerase, HIV-1 Reverse Transcriptase, other available DNA polymerases, reverse transcriptase, and other enzymes, including heat-stable enzymes, which will facilitate combination of the nucleotides in the proper manner to form the primer extension products which are complementary to each nucleic acid strand. Generally, the synthesis will be initiated at the 3' end of each primer and proceed in the 5' direction along the identifier strand, until synthesis terminates, producing molecules of different lengths. There may be inducing agents, however, which initiate synthesis at the 5' end and proceed in the above direction, using the same process as described above.

The inducing agent also may be a compound or system which will function to accomplish the synthesis of RNA primer extension products, including enzymes. In preferred embodiments, the inducing agent may be a DNA-dependent RNA polymerase such as T7 RNA polymerase, T3 RNA polymerase or SP6 RNA polymerase. These polymerases produce a complementary RNA polynucleotide. The high turnover rate of the RNA polymerase amplifies the starting polynucleotide as has been described by Chamberlin et al., The Enzymes, ed. P. Boyer, pp. 87-108, Academic Press, New York (1982). Amplification systems based on transcription have been described by Gingeras et al., in PCR Protocols, A Guide to Methods and Applications, pp. 245-252, Innis et al., eds, Academic Press, Inc., San Diego, Calif. (1990).

If the inducing agent is a DNA-dependent RNA polymerase and, therefore incorporates ribonucleotide triphosphates, sufficient amounts of ATP, CTP, GTP and UTP are admixed to the primer extension reaction admixture and the resulting solution is

treated as described above.

The newly synthesized strand and its complementary nucleic acid strand form a double-stranded molecule which can be used in the succeeding steps of the method. PCR amplification methods are described in detail in U.S. Pat. Nos. 4,683,192, 4,683,202, 4,800,159, and 4,965,188, and at least in several texts including PCR Technology: Principles and Applications for DNA Amplification, H. Erlich, ed., Stockton Press, New York (1989); and PCR Protocols: A Guide to Methods and Applications, Innis et al., eds., Academic Press, San Diego, Calif. (1990). The term "primer" as used herein refers to a polynucleotide whether purified from a nucleic acid restriction digest or produced synthetically, which is capable of acting as a point of initiation of nucleic acid synthesis when placed under conditions in which synthesis of a primer extension product which is complementary to a nucleic acid strand is induced, i.e., in the presence of nucleotides and an agent for polymerization such as DNA polymerase, reverse transcriptase and the like, and at a suitable temperature and pH. The primer is preferably single stranded for maximum efficiency, but may alternatively be in double stranded form. If double stranded, the primer is first treated to separate it from its complementary strand before being used to prepare extension products. Preferably, the primer is a polydeoxyribonucleotide. The primer must be sufficiently long to prime the synthesis of extension products in the presence of the agents for polymerization. The exact lengths of the primers will depend on many factors, including temperature and the source of primer. For example, depending on the complexity of the target sequence, a polynucleotide primer typically contains 10 to 25 or more nucleotides, although it can contain fewer nucleotides. Short primer molecules generally require cooler temperatures to form sufficiently stable hybrid complexes with identifier.

The primers used herein are selected to be "substantially" complementary to the different strands of each specific sequence to be synthesized or amplified. This means that the primer must be sufficiently complementary to non-randomly hybridize with its respective identifier strand. Therefore, the primer sequence may or may not reflect the exact sequence of the identifier. For example, a non-complementary nucleic acid can be attached to the 5' end of the primer, with the remainder of the primer sequence being substantially complementary to the strand. Such non-complementary fragments typically code for an endonuclease restriction site or used

as a linker to connect to a label, such as biotin.

Primers of the present invention may also contain a DNA-dependent RNA polymerase promoter sequence or its complement. See for example, Krieg et al., Nucl. Acids Res., 12:7057-70 (1984); Studier et al., J. Mol. Biol., 189:113-130 (1986); and Molecular Cloning: A Laboratory Manual, Second Edition, Maniatis et al., eds., Cold Spring Harbor, N.Y. (1989).

When a primer containing a DNA-dependent RNA polymerase promoter is used, the primer is hybridized to the polynucleotide strand to be amplified and the second polynucleotide strand of the DNA-dependent RNA polymerase promoter is completed using an inducing agent such as E. coli DNA polymerase I, or the Klenow fragment of E. coli DNA polymerase. The starting polynucleotide is amplified by alternating between the production of an RNA polynucleotide and DNA polynucleotide. This may be used for selective degradation of the RNA strand, which is prone to disintegration upon treatment with a strong base.

Primers may also contain a identifier sequence or replication initiation site for a RNA-directed RNA polymerase. Typical RNA-directed RNA polymerase include the QB replicase described by Lizardi et al., Biotechnology, 6:1197-1202 (1988). RNA-directed polymerases produce large numbers of RNA strands from a small number of identifier RNA strands that contain a identifier sequence or replication initiation site. These polymerases typically give a one million-fold amplification of the identifier strand as has been described by Kramer et al., J. Mol. Biol., 89:719-736 (1974).

In one embodiment, the present invention utilizes a set of polynucleotides that form primers having a priming region located at the 3'-terminus of the primer. The 3'-terminal priming portion of each primer is capable of acting as a primer to catalyze nucleic acid synthesis, i.e., initiate a primer extension reaction off its 3' terminus. One or both of the primers can additionally contain a 5'-terminal non-priming portion, i.e., a region that does not participate in hybridization to the preferred identifier. The 5'-part of the primer may be labelled as described herein above.

Brief Description of the Figures

Fig. 1a and 1b shows a picture of a gel resulting from example 1

Fig. 2 discloses a picture of a gel resulting from example 2.

Fig. 3 depicts a picture of gel which shows the results of example 3.

Fig. 4 discloses a schematic representation of two embodiments of the present invention.

Detailed Description of the drawings

Fig. 4 discloses details of the present invention. Initially, a library of complexes (frame 1) is provided. The potential drug candidate is represented by a circle and the identifier is indicated by a rectangle. The straight line between the identifier and the potential drug candidate illustrate a linkage which in some embodiments is cleavable and in other embodiments are durable under the conditions used. The potential drug candidate and the identifier are filled with a suitable pattern to illustrate that the individual complexes are different from one another.

In process step A, the library is subjected to a target in solution (frame 2). The part of the library which has affinity towards the target will bind to the target. The symbol β indicate a selectively cleavable linker, e.g. a PC spacer, cleavable by exposure to light of a certain wave length. Following process step I, the target may be immobilized to a solid support indicated by the container wall utilizing a linkage α (frame 3). The linkage α may be selectively cleavable or a durable linker. Preferably, in process step C, The non-binding part of the library is removed and the aggregate of the complex and the target is recovered (frame 4). The recovery of the aggregate of target and complex when in solution may be performed by size-exclusion chromatography.

After the recovery of the aggregate of complex and target, the target may be immobilized following process path I (frame 5). Preferably, using process step E, the linkage β is cleaved thus separating the potential drug candidate from the identifier (frame 6). After the separation, the identifier is isolated to decode the identity of the potential drug candidate. Prior to the isolation, the target bound to the potential drug candidate may be immobilized to facilitate the isolation. Optionally, the isolated identifier (frame 8) is amplified by e.g. PCR prior to the identification step.

Following the route starting with process step B, the library of complexes is subjected to an immobilized target. Complexes having affinity towards the target will

bind thereto (frame 3). In effect, the complex is immobilized. The non-binding complexes are washed away in process step D and the immobilized complex-target-aggregate is isolated (frame 5). In the event the linkage α is selectively cleavable, the process step I may be followed to form the isolated aggregate in frame 4. The identifier may be cleaved off and the decoded to identify the identity of the potential drug candidate. When the linkages α and β are both selectively cleavable, they are preferable orthogonal cleavable, i.e. the conditions which cleave the linkage α does not cleave linkage β and *visa versa*.

In the event the linker α is durable, i.e. not cleaved under the conditions used, the process step F is followed. The immobilized target is maintained connected to the solid support while the identifier is cleaved to the potential drug candidate (frame 7). The identifier is subsequently isolated using process step H (frame 8). Prior to decoding the identifier may be amplified by suitable process to obtain an appropriate amount of genetic material.

Examples

Example 1: Enrichment of biotin tagged DNA identifier.

100 pico mol of DNA identifier T1 of the sequence (GAGCGGATGCGTACATCTTGTACATGTCAATGCGATCGATCGACTGATCAAGATAGACCGAGTGCCG) was annealed to 50 pico mol of a biotin tagged primer F1-PC-B (Biotin-PC-CGGCACTCGGTCTATCTT). PC is a photo cleavable group and is obtainable as a phosphoramidite from Glen Research (Products cat# 10-4913). Subsequently the F1-PC-B primer was extended by DNA polymerase (Sequenase from Upstate Biotechnology Cat# 70775Y). Likewise 50 pico mol of a DNA identifier T2 (GAGCGGATGCGTACATCTACGATGGATGCTCCAGGTCGCAAGATAGACCGAGTGCCG) was annealed to 100 pico mol of a primer F1 (CGGCACTCGGTCTATCTT) and extended by DNA polymerase. This generated two pools of double stranded DNA sequences each containing 50 pico mol of the following complexes, one of length 67 nucleotides (termed T1+B) and one of length 57 nucleotides termed (T2-B).

Enrichment of the biotin tagged T1+B complex was performed in Reacti-Bind™ streptavidin coated clear strip plates (Pierce Biotechnology, Cat # 15120). The two

pools of double stranded DNA, T1+B and T2-B where mixed in ratios 1 pmol of T1+B complex and 1 pmol of T2-B complex, 1/10 pmol of T1+B complex and 1 pmol of T2-B complex, 1/100 pmol of T1+B complex and 1 pmol of T2-B complex, 1/1,000 pmol of T1+B complex and 1 pmol of T2-B complex, 1/10,000 pmol of T1+B complex and 1 pmol of T2-B complex, 1/100,000 pmol of T1+B complex and 1 pmol of T2-B complex and 1/1,000,000 pmol of T1+B complex and 1 pmol of T2-B complex. All where mixed in 100 µL buffer A (Tris buffered saline, 0.05% Tween 20, 1% Bovine serum albumin, 0.1 mg/mL herring sperm DNA). Incubation in streptavidin coated wells was done for 30 min at 25°C. After ligand binding wells were washed 30 times with 250 µL buffer A during one hour. Thereafter 100 µL buffer A was applied to each well and the wells where exposed to UV light at 350 nano meters for 30 seconds in order to cleave the PC spacer thereby releasing the T1 identifier from the biotin molecule. Following exposure to UV light the elution volume was removed immediately and analyzed for the presence of DNA strands T1 and T2 by polymerase chain reaction (PCR).

The content of T1 and T2 following enrichment for biotin binding molecules was analysed by generating ^{32}P phosphate labeled PCR products of T1 and T2. Both T1 and T2 PCR products where generated by the use of the two primers F1 (CGGCACTCGGTCTATCTT) and R1 (GAGCGGATGCGTACATCT). R1 was labeled with ^{32}P phosphate by the use of T4 polynucleotide kinase from Promega (Cat# M4101). PCR was performed by adding 10 µL of the eluted volume together with 15 µL of water containing 1 pico mole of both F1 and ^{32}P-labeled R1 to puReTaq™ Ready-To-Go™ PCR Beads (Amersham Biosciences, Cat# 407513-96). PCR was performed for 30 cycles by annealing at 58°C and extending at 72°C. PCR generated products where separated by urea containing polyacrylamide gel electrophoresis and visualized by radiography.

Eight wells precoated with streptavidin were used for the enrichment experiment shown in Fig. 1A. Lane 1 in Fig 1A. represents PCR amplification on eluate from a well that was incubated with Buffer A without any identifier complexes. Lane 2 was incubated with Buffer A containing 1 pmol of T1+B complex and 1 pmol of T2-B complex. Lane 3 was incubated with Buffer A containing 1/10 pmol of T1+B complex and 1 pmol of T2-B complex. Lane 4 was incubated with Buffer A containing 1/100 pmol of T1+B complex and 1 pmol of T2-B complex. Lane 5 was incubated with

Buffer A containing 1/1,000 pmol of T1+B complex and 1 pmol of T2-B complex. Lane 6 was incubated with Buffer A containing 1/10,000 pmol of T1+B complex and 1 pmol of T2-B complex. Lane 7 was incubated with Buffer A containing 1/100,000 pmol of T1+B complex and 1 pmol of T2-B complex. Lane 8 was incubated with Buffer A containing 1/1,000,000 pmol of T1+B complex and 1 pmol of T2-B complex.

Lane 9 in Fig. 1A represents PCR amplification of the input into well 2. Lane 10 represents PCR amplification of the input into well 3. Lane 11 represents PCR amplification of the input into well 4. Lane 12 represents PCR amplification of the input into well 5. Lane 13 represents PCR amplification of the input into well 6. Lane 14 represents PCR amplification of the input into well 7. Lane 15 represents PCR amplification of the input into well 8. Lane 16 represents PCR amplification of water (Blank).

In Fig. 1B the PCR products represented in Fig. 1A were subjected to additional 30 cycles of PCR amplification. This revealed T1 and T2 identifiers of very low abundance. Lane 1 represents further PCR amplification of PCR product as represented in Fig. 1B lane 1. Lane 2 represents further PCR amplification of PCR product as represented in Fig. 1B lane 2. Lane 3 represents further PCR amplification of PCR product as represented in fig 1 lane 3. Lane 4 represents further PCR amplification of PCR product as represented in Fig. 1B lane 4. Lane 5 represents further PCR amplification of PCR product as represented in Fig. 1B lane 5. Lane 6 represents further PCR amplification of PCR product as represented in Fig. 1B lane 6. Lane 7 represents further PCR amplification of PCR product as represented in Fig. 1 lane 7. Lane 8 represents further PCR amplification of PCR product as represented in Fig. 1B lane 8.

Lane 9 of Fig. 1B represents PCR amplification of water (Blank). Lane 10 represents PCR amplification of the input into well 2 of Fig. 1B. Lane 11 represents PCR amplification of the input into well 3 of Fig. 1B. Lane 12 represents PCR amplification of the input into well 4 of Fig. 1B. Lane 13 represents PCR amplification of the input into well 5 of Fig. 1B. Lane 14 represents PCR amplification of the input into well 6 of Fig. 1B. Lane 15 represents PCR amplification of the input into well 7 of Fig. 1B. Lane 16 represents PCR amplification of the input into well 8 of Fig. 1B. Lane 17 represents PCR amplification of water (Blank).

From Fig. 1A lane 6 it can be inferred that the double stranded DNA complex T1+B was enriched more than ten thousand fold over the T2-B complex. Further enrichment can not be deduced from Fig. 1A since the number of identifiers recovered and analyzed by PCR in lane 7 and 8 where to few to produce detectable bands after 30 rounds of PCR. When subjecting the PCR products shown in Fig. 1B to an additional 30 rounds of PCR the T2 product, in addition to the T1 product, appears in Fig. 1B lanes 5, 6, 7 and 8. The T2 product becomes the dominating PCR product in lane 8 suggesting that enrichment of the T1+B complex over the T2-B complex is between one hundred thousand and one million fold.

Example 2: Enrichment of di-nitro-phenol tagged DNA identifier.

100 pico mol of a DNA identifier, T1 (GAGCGGATGCGTACATCTTGTACATGTCA-ATGCGATCGATCGACTGATCAAGATAGACCGAGTGCCG) was annealed to 50 pico mol of a di-nitrophenol (DNP) tagged primer F1-PC-DNP (DNP-PC-CGGCACTCGGTCTATCTT). PC is a photo cleavable group obtainable as a phosphoramidite from Glen Research (Products cat# 10-4913). Subsequently the F1-PC-DNP primer was extended by DNA polymerase (Sequenase from Upstate Biotechnology Cat# 70775Y). Likewise 50 pico mol of a DNA identifier T2 (GAGCGGATGCGTACATCTACGATGGATGCTCCAGGTCGCAAGATAGAC-CGAGTGCCG) was annealed to 100 pico mol of a primer F1 (CGGCACTCGGTCTATCTT) and extended by DNA polymerase. This generated two pools of double stranded DNA sequences each containing 50 pico mol of the following complexes, one of length 67 nucleotides termed T1+DNP and one of length 57 nucleotides termed T2-DNP.

Enrichment of the DNP tagged T1+DNP complex was performed by coating 1 µg anti DNP antibody in Nunc Immunomodule U8 Maxisorp (Biotecline cat # nun-47507). The two pools of double stranded DNA, T1+DNP and T2-DNP where mixed in ratios 1 pmol of T1+DNP complex and 1 pmol of T2-DNP complex, 1/10 pmol of T1+ DNP complex and 1 pmol of T2-DNP complex, 1/100 pmol of T1+DNP complex and 1 pmol of T2-DNP complex, 1/1000 pmol of T1+DNP complex and 1 pmol of T2-DNP complex, 1/10000 pmol of T1+DNP complex and 1 pmol of T2-DNP complex, 1/100000 pmol of T1+DNP complex and 1 pmol of T2-DNP complex and 1/1000000 pmol of T1+DNP complex and 1 pmol of T2-DNP complex. All where mixed in 100 µL buffer A (Tris buffered saline, 0.05% Tween 20, 1% Bovine serum

albumin, 0.1 mg/mL herring sperm DNA). Incubation in anti-DNP antibody coated wells was done for 30 min at 25°C. After ligand binding all wells were washed 30 times with 250 µL buffer A during one hour. Thereafter 100 µL buffer A was applied to each well and the wells where exposed to UV light at 350 nano meters for 30 seconds in order to cleave the PC spacer thereby releasing the T1 identifier from the DNP molecule. Following exposure to UV light the elution volume was removed immediately and analysis for the presence of DNA strands T1 and T2 by polymerase chain reaction (PCR).

The content of T1 and T2 following enrichment for DNP containing molecules was analysed by generating ^{32}P phosphate labeled PCR products of T1 and T2. Both T1 and T2 PCR products where generated by the use of the two primers F1 (CGGCACTCGGTCTATCTT) and R1 (GAGCGGATGCGTACATCT). R1 was labeled with ^{32}P phosphate by the use of T4 polynucleotide kinase from Promega (Cat# M4101). PCR was performed by adding 10 µL of the eluted volume together with 15 µL of water containing 1 pico mole of both F1 and ^{32}P-labeled R1 to puReTaq™ Ready-To-Go™ PCR Beads (Amersham Biosciences, Cat# 407513-96). PCR was performed for 30 cycles by annealing at 58°C and extending at 72°C. The PCR products where diluted 100 fold and subjected to an additional 30 rounds of PCR amplification as described above. PCR generated products where separated by urea containing polyacrylamide gel electrophoresis and visualized by radiography.

Eight wells coated with rabbit anti DNP antibody, 1 µg/well (DAKOCytomation # V0401) were used for the enrichment experiment shown in Fig. 2. Lane 1 in fig 2. represents PCR amplification of elutate from a well that was incubated with Buffer A without any identifier complexes. Lane 2 is PCR amplification of identifiers eluted from a well that was incubated with Buffer A containing 1 pmol of T1+DNP complex and 1 pmol of T2-DNP complex. Lane 3 is 1/10 pmol of T1+ DNP complex and 1 pmol of T2-DNP complex. Lane 4 is 1/100 pmol of T1+DNP complex and 1 pmol of T2-DNP complex. Lane 5 is 1/1000 pmol of T1+DNP complex and 1 pmol of T2-DNP complex. Lane 6 is 1/10000 pmol of T1+DNP complex and 1 pmol of T2-DNP complex. Lane 7 is 1/100000 pmol of T1+DNP complex and 1 pmol of T2-DNP complex. Lane 8 is 1/1000000 pmol of T1+DNP complex and 1 pmol of T2-DNP complex. Lane 9 of fig 1. represents PCR product of water for control. Lane 10 represents PCR amplification of the input into the well that was subjected to selection

and analyzed by PCR in lane 2 of fig 2. Lane 11 represents PCR amplification of the input into the well analyzed in lane 3 of fig 2. Lane 12 represents PCR amplification of the input into the well analyzed in lane 4 of fig. 2. Lane 13 represents PCR amplification of the input into the well analyzed in lane 5 of fig. 1. Lane 14 represents PCR amplification of the input into the well analyzed in lane 6 of Fig. 2. Lane 15 represents PCR amplification of the input into the well analyzed in lane 7 of Fig 2. Lane 16 represents PCR amplification of the input into the well analyzed in lane 8 of Fig 2. Lane 17 represents PCR amplification of water for control.

From Fig. 2 lane 8 it can be concluded that the double stranded DNA complex T1+DNP was enriched more than one million fold over the T2-DNP complex. A weak T2-DNP band is appearing in lane 8 suggesting that enrichment of T1+DNP is considerably higher than one million fold. Products from both T1 and T2 in lane 1 appear despite no input of T1+DNP or T2-DNP complexes in this lane. This represent cross contamination from handling of identifiers resulting in weak product bands after two by 30 cycles of PCR. PCR product bands from T1 disappear in input lanes 13 to 16 since the T2 identifier is much in excess and take over the PCR reaction utilizing the 1 pico mole of both F1 and ^{32}P-labeled R1 primers. This is also seen in the bottom of the gel where the ^{32}P-labeled R1 primer is disappearing.

Example 3: Enrichment of di-nitro-phenol tagged DNA identifier in the presence of a biotin tagged identifier. (A library of two displayed molecules)

100 pico mol of a DNA identifier, T1 (GAGCGGATGCGTACATCTTGTACATGTCAATGCGATCGATCGACTGATCAAGATAGACCGAGTGCCG) was annealed to 50 pico mol of a biotin tagged primer F1-PC-B (Biotin-PC-CGGCACTCGGTCTATCTT). PC is a photo cleavable group obtainable as a phosphoramidite from Glen Research (Products cat# 10-4913). Subsequently the F1-PC-B primer was extended by DNA polymerase. (Sequenase from Upstate Biotechnology Cat# 70775Y).

Likewise, 100 pmol T2 (GAGCGGATGCGTACATCTACGATGGATGCTCCAGGTCGCAAGATAGACCGAGTGCCG) was annealed to 50 pico mol of a di-nitrophenol (DNP) tagged primer F1-PC-DNP (DNP-PC-CGGCACTCGGTCTATCTT). Subsequently, the F1-PC-DNP primer was extended by DNA polymerase as described above.

This generated two pools of double stranded DNA sequences respectively containing 50 pico mol of the following complexes, one of length 67 nucleotides termed T1+B and one of length 57 nucleotides termed T2+DNP.

Enrichment of the DNP tagged T2+DNP complex was performed by coating 1 µg/well anti DNP antibody (DAKOCytomation # V0401) in Nunc Immunomodule U8 Maxisorp (Biotecline cat # nun-47507). The two pools of double stranded DNA, T2+DNP and T1+B where mixed in ratios 1 pmol of T2+DNP complex and 1 pmol of T1+B complex, 1/10 pmol of T2+DNP complex and 1 pmol of T2+B complex, 1/100 pmol of T2+DNP complex and 1 pmol of T1+B complex, 1/1000 pmol of T2+DNP complex and 1 pmol of T1+B complex, 1/10000 pmol of T2+DNP complex and 1 pmol of T1+B complex, 1/100000 pmol of T2+DNP complex and 1 pmol of T1+B complex and 1/1000000 pmol of T2+DNP complex and 1 pmol of T1+B complex. All where mixed in 100 µL buffer A (Tris buffered saline, 0.05% Tween 20, 1% Bovine serum albumin, 0.1 mg/mL herring sperm DNA). Incubation in anti-DNP antibody coated wells was done for 90 min at 25°C. After ligand binding all wells were washed 30 times with 250 µL buffer A during one hour. Thereafter 100 µL buffer A was applied to each well and the wells where exposed to UV light at 350 nano meters for 30 seconds in order to cleave the PC spacer thereby releasing the T2 identifier from the DNP molecule. Following exposure to UV light the elution volume was removed immediately and analysis for the presence of DNA strands T1 and T2 by polymerase chain reaction (PCR).

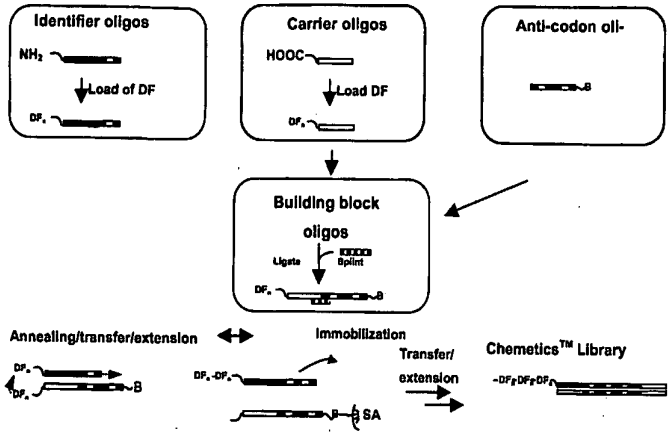
The content of T1 and T2 following enrichment for DNP containing molecules was analysed by generating ^{32}P phosphate labeled PCR products of T1 and T2. Both T1 and T2 PCR products where generated by the use of the two primers F1 (CGGCACTCGGTCTATCTT) and R1 (GAGCGGATGCGTACATCT). R1 was labeled with ^{32}P phosphate by the use of T4 polynucleotide kinase from Promega (Cat# M4101). PCR was performed by adding 10 µL of the eluted volume together with 15 µL of water containing 1 pico mole of both F1 and ^{32}P-labeled R1 to puReTaq™ Ready-To-Go™ PCR Beads (Amersham Biosciences, Cat# 407513-96). PCR was performed for 30 cycles by annealing at 58°C and extending at 72°C. PCR generated products where separated by urea containing polyacrylamide gel electrophoresis and visualized by radiography.

Eight wells coated with rabbit anti DNP antibody, 1 µg/well were used for the enrichment experiment shown in Fig. 3. Lane 1 in fig 3. represents PCR amplification of elutate from a well that was incubated with Buffer A without any identifier complexes. Lane 2 is PCR amplification of identifiers eluted from a well that was incubated with Buffer A containing 1 pmol of T2+DNP complex and 1 pmol of T1+B complex. Lane 3 is 1/10 pmol of T2+DNP complex and 1 pmol of T1+B complex. Lane 4 is 1/100 pmol of T2+DNP complex and 1 pmol of T1+B complex. Lane 5 is 1/1000 pmol of T2+DNP complex and 1 pmol of T1+B complex. Lane 6 is 1/10000 pmol of T2+DNP complex and 1 pmol of T1+B complex. Lane 7 is 1/100000 pmol of T2+DNP complex and 1 pmol of T1+B complex. Lane 8 is 1/1000000 pmol of T2+DNP complex and 1 pmol of T1+B complex. Lane 9 of fig 1. represents PCR product of water for control. Lane 10 represents PCR amplification of the input into the well that was subjected to selection and analyzed by PCR in lane 2 of Fig. 3. Lane 11 represents PCR amplification of the input into the well analyzed in lane 3 of Fig. 1. Lane 12 represents PCR amplification of the input into the well analyzed in lane 4 of Fig. 3. Lane 13 represents PCR amplification of the input into the well analyzed in lane 5 of Fig 3. Lane 14 represents PCR amplification of the input into the well analyzed in lane 6 of Fig 3. Lane 15 represents PCR amplification of the input into the well analyzed in lane 7 of Fig 3. Lane 16 represents PCR amplification of the input into the well analyzed in lane 8 of Fig. 3. Lane 17 represents PCR amplification of water for control.

From Fig. 3 lane 5 it can be concluded that the double stranded DNA complex T2+DNP was enriched more than one thousand fold over the T2-B complex. Only weak T2-DNP bands are appearing in lane 3-5 suggesting that enrichment of T2+DNP is considerably higher than one thousand fold. Further fold enrichment can not be concluded from this experiment. Additional rounds of PCR could possibly reveal the ratio between T1 and T2 in lanes 6-8. PCR product bands from T2 disappear in input lanes 13 to 16 since the T1 identifier is much in excess and take over the PCR reaction utilizing the 1 pico mole of both F1 and ^{32}P-labeled R1 primers. This is also seen in the bottom of the gel where the ^{32}P-labeled R1 primer is disappearing.

Example 4: Selection of an integrin αVβ3 ligand from a 484-member small molecule library encoded by chemetics™.

Overview of the procedure



DF : Drug fragment / chemical entity
B : Biotin
SA: Streptavidin

The method for producing a library of bifunctional complexes, in which each member of the library comprises a synthetic molecule and an identifier that may be decoded to establish the synthetic history of the synthetic molecule comprises several steps, exemplified below. In a first step (General procedure 1), four different identifier oligonucleotides are loaded with a scaffold molecule or drug fragment. In this example the loading is conducted using an amino group on the identifier oligo as the attachment point for the drug fragment/ scaffold molecule. The identifiers may be regarded as the nascent bifunctional complexes.

To prepare the building block oligos, identical carrier oligos are initially loaded with eleven different drug fragments using general procedure 2. The eleven loaded carrier oligos are then ligated to anti-codon oligos of the first and the second round using general procedure 3, thereby obtaining 11 building blocks for the first round and eleven building blocks for the second round.

The library formation is described in detail in general procedure 4 and includes the mixing of the four different identifier oligos with the eleven different building blocks of the fist round. To bias the library one of the identifiers and one of the first round building blocks were added in an amount 100 below the amount of the other components. At conditions providing for annealing between the identifiers and the building blocks, a cross-link between the scaffold molecules of the identifier oligo and the drug fragments were effected. The identifier oligos were then extended using a polymerase and using the anti-codon of the building block as the identifier. After the extension, the drug fragment is released from the building block by cleavage of a linkage between the drug fragment and the oligo. The spent building block oligo is removed by streptavidin beads.

The second round includes the addition of building blocks to the nascent identifier-synthetic molecule complex obtained in the first round. To bias the library, one of the eleven second round building blocks was added in an amount 100 times below the amount used for the 10 other building blocks. The second round follows the same scheme as depicted above for the first round. The library formed is of 4 * 11 * 11 = 484 members. One of the members, which is a known ligand for the target, appears only in a concentration of the library of one out of $3 * 10^8$ bifunctional complexes.

The library is then subjected to a selection process, as disclosed in general procedure 5. The selection involves addition of the library to wells coated with immobilized target. After incubation of the library with the target, non-binding members of the library is removed by washing and a linkage between the synthetic molecule and the indentifier is cleaved. The cleaved off identifiers were collected and amplified by PCR. The amplified identifiers were decoded using general procedure 6.

General procedure 1: Loading of identifier oligos

10 µL triethanolamine (TEA) (0.1 M in DMF) was mixed with 10 µL Building Block (BB) with Pent-4-enal as an amine protection group (0.1 M in DMSO). From this mixture 6.7 µL was taken and mixed with 3.3 µL EDC [1-Ethyl-3-(3-Dimethylaminopropyl) carbodiimide Hydrochloride] (0.1 M in DMF) and incubated 30 minutes at 25°C. 10 µL of the Building block-EDC-TEA mixture was added to 10 µL of amino oligo in 0.1 M HEPES buffer ((4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid, SIGMA), pH 7.5 and incubated with the oligo for 30 minutes.

During this half hour, another 6.7 µL of BB-TEA mix was mixed with3.3 µL EDC (0.1 M in DMF) and incubate for 30 minutes at 25°C. 10 µL of this second BB-EDC-TEA mixture was then added to the amino oligo mixture together with 10 µL of 0.1 M HEPES buffer to maintain a 1:1 ratio of DMSO/DMF: H_2O. Then the mixture was incubated for 30 minutes.

During this half hour, another 6.7 µL of BB-TEA mix was mixed with3.3 µL EDC (0.1 M in DMF) and incubate for 30 minutes at 25°C. 10 µL of this third BB-EDC-TEA mixture was then added to the amino oligo mixture together with 10 µL of 0.1 M HEPES buffer to maintain a 1:1 ratio of DMSO/DMF: H_2O. Then the mixture was incubated for 30 minutes.

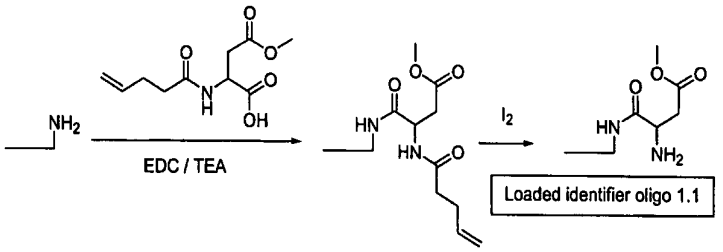
The loaded oligo was then purified by gel filtration with columns (Biospin P-6, Bio-Rad) equilibrated with water. The pent-4-enal amine protection group was then removed by addition of 0.25 volumes 25 mM I_2 in 1:1 water:tetrahydrofuran (THF) and incubation at 37°C for 2 hours. The mixture was then purified by gel filtration with spin columns (Biospin P-6, BioRad) equilibrated with water. Loaded identifier oligos were analyzed by ES-MS.

Example 4.1.1

Identifier oligo 1.1: 5'-
NSPACCTCAGCTGTGTATCGAGCGGCAGCGTTATCGTCG-3'

Sequence identifying the loaded fragment

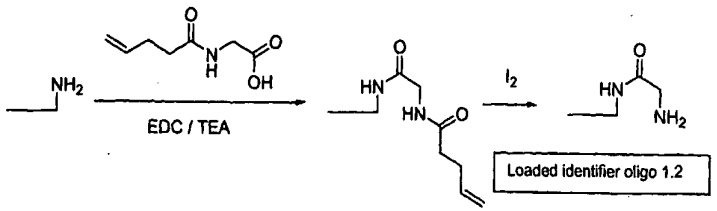
N : 5'-Amino-Modifier 5 (Glen research cat# 10-1905-90)
S : Spacer C3 CPG (Glen research cat# 20-2913-01)
P : PC Spacer Phosphoramidite (Glen research cat# 10-4913-90)



Loaded identifier oligo 1.1 analyzed by ES-MS:
Expected Mass : 11709 Da
Observed Mass : 11708 Da

Example 4.1.2
Identifier oligo 1.2: 5'- **NSP**ACCTCAGCTGTGTATCGAGCGGCAGCAGTGC-CGTCG-3'

N : 5'-Amino-Modifier 5 (Glen research cat# 10-1905-90)
S : Spacer C3 CPG (Glen research cat# 20-2913-01)
P : PC Spacer Phosphoramidite (Glen research cat# 10-4913-90)

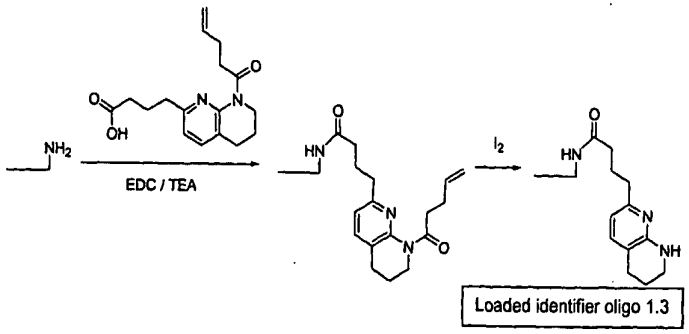


Loaded identifier oligo 1.2 analyzed by ES-MS:
Expected Mass : 11647 Da
Observed Mass : 11641 Da

Example 4.1.3

Identifier oligo 1.3: 5'- **NSP**ACCTCAGCTGTGTATCGAGCGGCAGCGCA-CACGTCG-3'

N : 5'-Amino-Modifier 5 (Glen research cat# 10-1905-90)
S : Spacer C3 CPG (Glen research cat# 20-2913-01)
P : PC Spacer Phosphoramidite (Glen research cat# 10-4913-90)



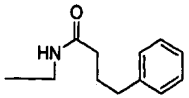
Loaded identifier oligo 1.2 analyzed by ES-MS:
Expected Mass : 11761 Da
Observed Mass : 11759 Da

Example 4.1.4

Identifier oligo 1.4: 5'- **NSP**ACCTCAGCTGTGTATCGAGCGGCAGCGGA-TACGTCG-3'

N : 5'-Amino-Modifier 5 (Glen research cat# 10-1905-90)
S : Spacer C3 CPG (Glen research cat# 20-2913-01)
P : PC Spacer Phosphoramidite (Glen research cat# 10-4913-90)

Loaded identifier oligo:



Expected Mass : 11775 Da
Observed Mass : 11775 Da

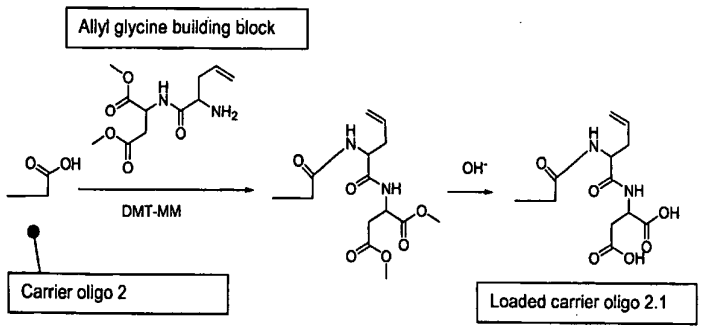
General procedure 2: Loading of carrier oligo

10-15 nmol of carrier oligo 2 was lyophilized and redissolved in 27.5 µl H_2O. To this was added 7.5 µl 1 M HEPES pH 7.5, 10 µl of 2-amino-pent-4-enal protected (allyl-glycine) building block (0.1 M in dimethyl sulfoxide), and 5 µl DMT-MM [4-(4,6-dimethoxy-1,3,5-thiazin-2-yl)-4-methylmorpholinium chloride] (0.5 M in water). The mixture was incubated 4-16 hours at 25-30°C. The oligo was purified by gel filtration (Biospin P-6, BioRad). To convert the methyl ester moiety of the building block to a carboxylic acid, 5 µl 0.4 M NaOH was added and the mixture was incubated 20 min at 80°C. The mixture was then neutralized by adding 10 µl 0.5 M HEPES pH 7.5 and 5 µl 0.4 M HCl. The loaded building block oligo was purified by gel filtration (Biospin P-6, BioRad) and analyzed by ES-MS

Carrier oligo 2: 3'-**2**GGAGTCGACACATAGCTCGC**p**-5'

2: Carboxy dT (Glen research cat# 10-1035-90)
p: 5' phosphate

Example 4.2.1

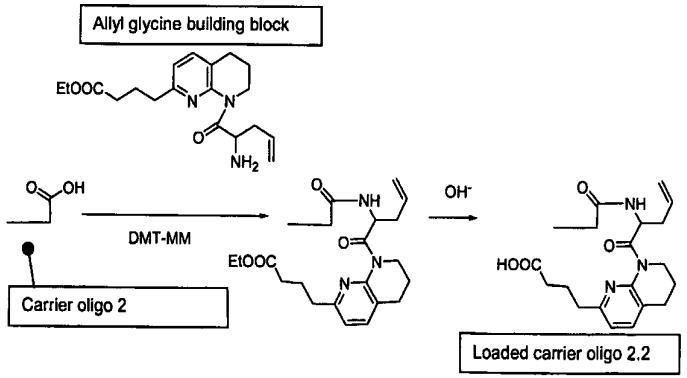


Loaded carrier oligo 2.1 analyzed by ES-MS:

Expected Mass : 6856 Da

Observed Mass : 6857 Da

Example 4.2.2

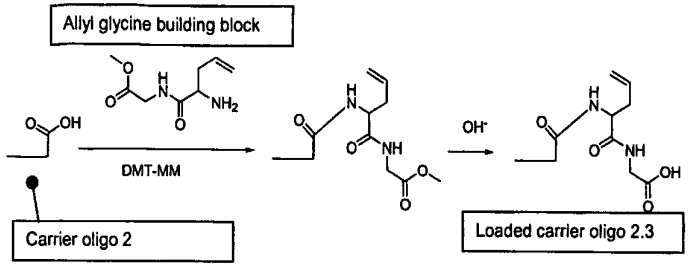


Loaded carrier oligo 2.2 analyzed by ES-MS:

Expected Mass : 6944 Da

Observed Mass : 6945 Da

Example 4.2.3

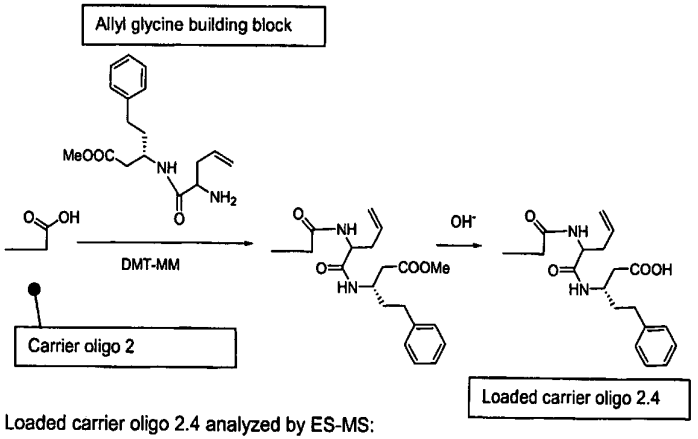


Loaded carrier oligo 2.3 analyzed by ES-MS:

Expected Mass : 6798 Da

Observed Mass : 6800 Da

Example 4.2.4

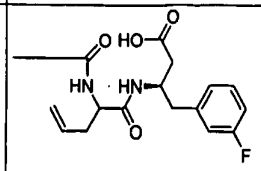
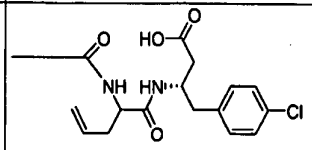
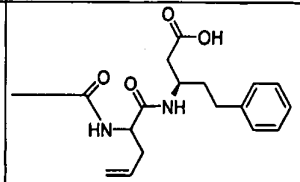
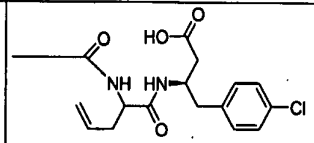
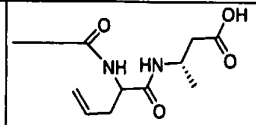
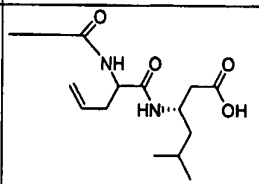


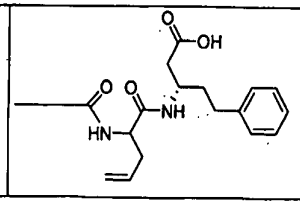
Loaded carrier oligo 2.4 analyzed by ES-MS:

Expected Mass : 6917 Da

Observed Mass : 6919 Da

Table I

Carrier oligo Example	Structure of loaded Carrier oligo	Expected Mass	Observed Mass
4.2.5		6924	6923
4.2.6		6940	6939
4.2.7		6920	6919
4.2.8		6940	6939
4.2.9		6830	6829
4.2.10		6871	6871

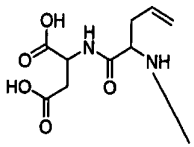
4.2.11		6920	6919

General procedure 3: ligation of anti-codon oligo with loaded carrier oligo

500 pmol loaded carrier oligo was mixed with 750 pmol anti-codon oligo and 750 pmol splint oligo. The mixture was lyophilized and redissolved in 15 µl water. Oligos were annealed by heating and slowly cooling to 20°C. 15 µl TaKaRa ligase mixture (Takara Bio Inc) was added and the reaction was incubated at 20°C for 1 hour. The mixture was purified by gel filtration (Biospin P-6, BioRad) and the efficiency of the ligation was checked by running an aliquot on a Novex TBE-UREA gel (Invitrogen).

Examples of building block oligos for first round of encoding

Example 4.3.1.1



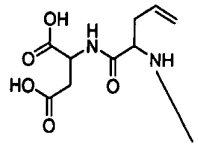
Loaded carrier Oligo 2.1

Anti-codon oligo

3'-2GGAGTCGACACATAGCTCGCp CGTCGIIIIIGCAGCCA-ATAGTCGT-X

::::::: :::::::

Splint oligo: TCGAGCG--GCAGCCA



Ligase

3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIG-CAGCCAATAGTCGT-X

Building block oligo 3.1.1

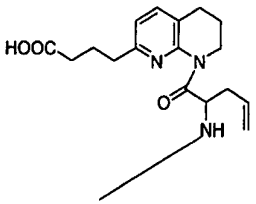
2: Carboxy dT (Glen research cat# 10-1035-90)

P: 5' phosphate

X: 5' biotin

Efficiency of ligation : > 95 %

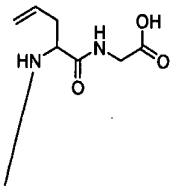
Example 4.3.1.2



3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCCGTGTGTCGT-X

Efficiency of ligation : > 95 %

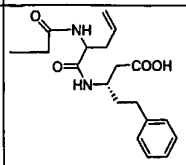
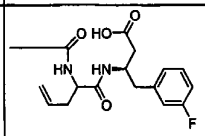
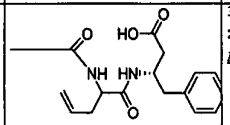
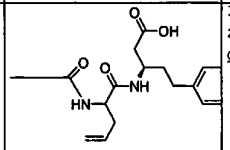
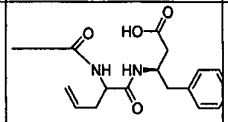
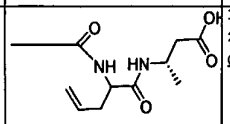
Example 4.3.1.3

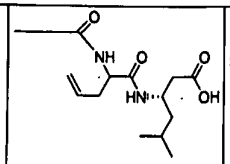
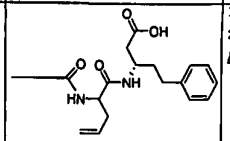


3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCTCACGGTCGT-X

Efficiency of ligation : > 95 %

Table II

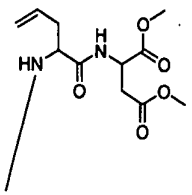
Building block oligo example	Structure of loaded Drug fragment	Building block oligo sequence 2: Carboxy dT (Glen research cat# 10-1035-90) X: 5' biotin	Ligation effi-ciency
4.3.1.4		3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCCCTATGTCGT-X	> 95 %
4.3.1.5		3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCGCGACGTCGT-X	> 95 %
4.3.1.6		3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCGACCAGTCGT-X	> 95 %
4.3.1.7		3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCACAAGGTCGT-X	> 95 %
4.3.1.8		3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCTGGACGTCGT-X	> 95 %
4.3.1.9		3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCGCTCGGTCGT-X	> 95 %

4.3.1.10		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCCATA GGTCGT-X	> 95 %
4.3.1.11		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCCCGG AGTCGT-X	> 95 %

Examples of building block oligos for second round of encoding

Example 4.3.2.1

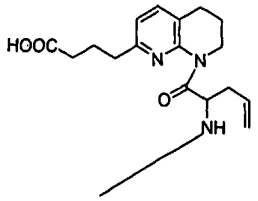
Building block oligo 3.2.1 :



3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGTCAATA-
CAGCTTAGACGGTAGATTTX

Efficiency of ligation : > 95 %

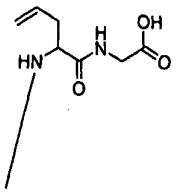
Example 4.3.2.2



3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGTCGTGTCAG-
CTTAGACGGTAGATTTX

Efficiency of ligation : > 95 %

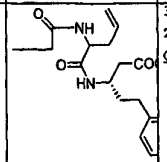
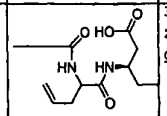
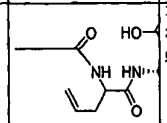
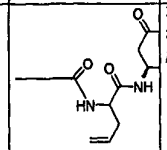
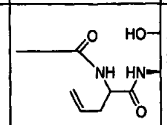
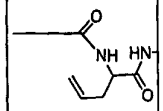
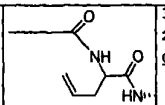
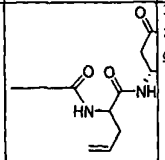
Example 4.3.2.3



3'-2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGTTCACG-
CAGCTTAGA-CGGTAGATTTX

Efficiency of ligation : > 95 %

Table III

Building block oligo example	Structure of loaded Drug fragment	Building block oligo sequence 2: Carboxy dT (Glen research cat# 10-1035-90) X: 5' biotin	Ligation effi-ciency
4.3.2.4		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGT CCTATCAGCTTAGACGGTAGATTTX	> 95 %
4.3.2.5		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGT GCGACCAGCTTAGACGGTAGATTTX	> 95 %
4.3.2.6		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGT GACCACAGCTTAGACGGTAGATTTX	> 95 %
4.3.2.7		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGT ACAAGCAGCTTAGACGGTAGATTTX	> 95 %
4.3.2.8		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGT TGGACCAGCTTAGACGGTAGATTTX	> 95 %
4.3.2.9		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGT GCTCGCAGCTTAGACGGTAGATTTX	> 95 %
4.3.2.10		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGT CATAGCAGCTTAGACGGTAGATTTX	> 95 %
4.3.2.11		3'- 2GGAGTCGACACATAGCTCGCCGTCGIIIIIGCAGCIIIIIGTCGT CCGGACAGCTTAGACGGTAGATTTX	> 95 %

General procedure 4: Encoding a small molecule library by chemetics™

Example 4.4.1: Encoding a 484-member small molecule library by chemetics™

Example 4.4.1.1 First encoding round

2 pmol of loaded identifier oligo 1.1 was combined with 200 pmol of each loaded identifier oligo 1.2, 1.3, and 1.4. (602 pmol loaded identifier oligos in total). These were mixed with 0.7 pmol building block oligo 3.1.3., and 72.7 pmol each of 10 different other first round building block oligos (eg. 3.1.1 and 3.1.2; 727 pmol loaded building block oligos in total). The oligos were lyophilized and redissolved in 50 µl extension buffer (EX) [20 mM HEPES, 150 mM NaCl, 8 mM $MgCl_2$]. The mixture was heated to 80°C and slowly cooled to 20°C to allow efficient annealing of identifier and building block oligos. 5 µl of 0.5 M DMT-MM in water was added and the mixture was incubated at 37°C for 4 hours.

Extension of the identifier oligo on the building block oligo identifier was performed by adding 3 µl of a 10 mM mixture of each deoxynucleotide triphosphate [dATP, dGTP, dCTP, dTTP] and 3 µl of 13 units/µl Sequenase (Amersham Biosciences). The mixture was subsequently incubated at 30°C overnight. Then 3 µl of 2M NaOH was added and the mixture was incubated for 80°C for 10 minutes followed by neutralization by addition of 3 µl 2M HCl. The mixture was then purified by passing through a gel filtration column (Biospin P-6, BioRad). 0.25 volumes of 25 mM I_2 in

1:1 THF:water was added, mixed and incubated at 37°C for 2 hours. 60 µl binding buffer (BF) [100 mM HEPES, 150 mM NaCl] and water ad 300 µl was added.

The mixture was added to streptavidin-sepharose beads (Amersham Biosciences) pre-washed 3 times in BF buffer and incubated at room temperature for 10 minutes followed by incubation on ice for 10 minutes with gentle stirring. The beads were then washed three times with water. Extended identifier oligos were stripped from the building block oligos bound to the streptaviding-sepharose beads by applying 100 µl NH3 1:1 in water and incubating at room temperature for 5 minutes.

4.4.1.2 Second encoding round

To the eluate was added 0.36 pmol second round loaded building block oligo 3.2.2 and 36.4 pmol each of 10 different other second round building block oligos (eg. 3.2.1 and 3.2.3; 364 pmol loaded second round building block oligos in total) and the mixture was lyophilized and redissolved in 50 µl EX buffer. The encoding was performed essentially as described under 4.1.1.

4.4.1.3 Final extension

The eluted identifier oligo were lyophilized and dissolved in 50 µl EX buffer. Then 200 pmol primer E38 [5'-XTTTTAGATGGCAGAT-3', X=CXS Biotin] was added. Annealing was performed by heating the mixture to 80°C and slowly cooling to 20°C. Extension of the identifier oligo was performed by adding 3 µl of a 10 mM mixture of each deoxynucleotide triphosphate [dATP, dGTP, dCTP, dTTP] and 3 µl of 13 units/µl Sequenase. The mixture was subsequently incubated at 30°C for 2 hours. The mixture was then purified by passing through a gel filtration column (Biospin P-6, BioRad). This eluated was used for selection. An aliquot (sample 4.1.3) was removed for analysis of the inpout in the selection procedure.

General procedure 5: selection

Maxisorp ELISA wells (NUNC A/S, Denmark) were coated with each 100 µL 2µg/mL integrin αVβ3 (Bachem) in PBS buffer [2.8 mM NaH_2PO_4, 7.2 mM Na_2HPO_4, 0.15 M NaCl, pH 7.2] overnight at 4°C. Then the integrin solution was substituted for 200 µl blocking buffer [TBS, 0.05% Tween 20 (Sigma P-9416), 1% bovine serum albumin (Sigma A-7030), 1 mM $MnCl_2$] which was left on for 3 hours at room temperature. Then the wells were washed 10 times with blocking buffer and the encoded library

was added to the wells after diluting it 100 times with blocking buffer. Following 2 hours incubation at room temperature the wells were washed 10 times with blocking buffer. After the final wash the wells were cleared of wash buffer and subsequently inverted and exposed to UV light at 300-350 nm for 30 seconds. Then 100 µl blocking buffer without Tween-20 was immediately added to each well, the wells were shaken for 30 seconds, and the solutions containing eluted identifiers were removed for PCR analysis (sample 5.1)

General procedure 6: analysis of selection input and output

PCR was performed on the input for (sample 4.3.1) and output of (sample 5.1) the selection using primers corresponding to the 5' end of the identifier oligos and the E38 primer. PCR was performed using Ready-To-Go (RTG) PCR beads (Amersham Biosciences) and 10 pmol each primer in a reaction volume of 25 µl. The PCR reaction consisted of an initial denaturation step of 94°C for 2 minutes followed by 30-45 cycles of denaturation at 94°C for 30 seconds, annealing at 58°C for 1 minute and extension at 72°C for 1 minute. A final extension step of 2 minutes at 72°C was included. The PCR products were resolved by agarose gel electrophoresis and the band corresponding to the expected size was cut from the gel and purified using QIAquick Gel Extraction Kit (QIAGEN).

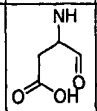
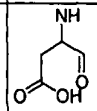
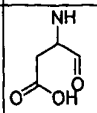
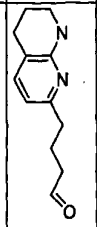
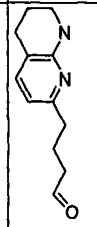
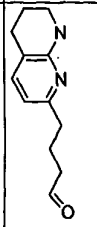
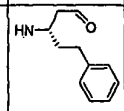
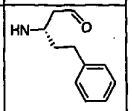
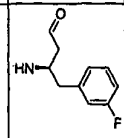
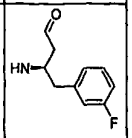
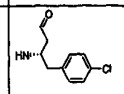
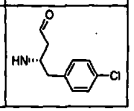
To sequence individual PCR fragments the purified PCR products were cloned into the pCR4-TOPO vector (Invitrogen) according to the manufacturer's instructions. The resulting mixture was used for transformation of TOP10 *E. coli* cells (Invitrogen) using standard procedures. The cells were plated on growth medium containing 100 µg/ml ampicillin and left at 37°C for 12-16 hours.

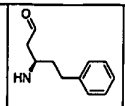
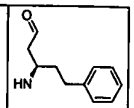
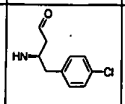
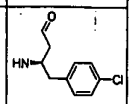
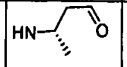
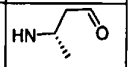
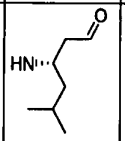
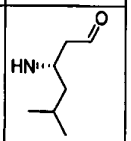
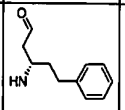
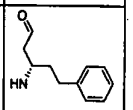
Individual *E.coli* clones were picked and transferred to PCR wells containg 50 µl water. These wells were then boiled for 5 minutes and 20 µl mixture from each well was used in a PCR reaction using RTG PCR beads and 5 pmol each of M13 forward and reverse primers according to the manufacturer's instructions. A sample of each PCR product was then treated with Exonuclease I (USB) and Shrimp Alkaline Phosphatase (USB) to remove degrade single stranded DNA and dNTPs and sequenced using the DYEnamic ET cycle sequencing kit (Amersham Biosciences) according to the manufacturer's instructions and the reactions were analyzed on a

MegaBace 4000 capillary sequencer (Amersham Biosciences). Sequence outputs were analyzed with ContigExpress software (Informax Inc.).

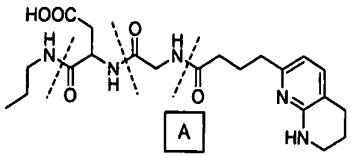
Overview of drug fragments present in the library:

Table IV

Identifier			Building block oligo for first round			Building block oligo for second round		
Oligo	Relative amount in library	Structure of drug fragment	Oligo	Relative amount in library	Structure of transferred drug fragment	Oligo	Relative amount in library	Structure of transferred drug fragment
1.1	100		3.1.1	1		3.2.1	100	
1.2	1		3.1.2	100		3.2.2	100	
1.3	100		3.1.3	100		3.2.3	1	
1.4	100		3.1.4	100		3.2.4	100	
			3.1.5	100		3.2.5	100	
			3.1.6	100		3.2.6	100	

			3.1.7	100		3.2.7	100	
			3.1.8	100		3.2.8	100	
			3.1.9	100		3.2.9	100	
			3.1.10	100		3.2.10	100	
			3.1.11	100		3.2.11	100	

The library had the potential to encode the integrin αVβ3 ligand A (Molecule **7** in Feuston B. P. et al., Journal of Medicinal Chemistry **2002**, *45*, 5640-5648) from 1 out of $3*10^8$ identifiers.



As can be seen from the table above, the library had the potential to encode ligand A for every $3*10^6$ identifiers (1 x 1 x 1 = 1 out of every 301 x 1001 x 1001 ~ $3*10^8$)

Example 4.6.1: Result of sequencing analysis of input for selection procedure and output from selection procedure.

The codon combination compatible with encoding of ligand A was not found in 28 sequences derived from the encoded library before selection in agreement with the expected low abundance of this codon combination (1 in $3*10^8$).

A codon combination compatible with encoding of ligand A was found in 5 out of 19 sequences derived from the encoded library after selection in integrin αVβ3-coated wells.
These numbers correspond to an enrichment factor of $(3*10^8 / (19 / 7)) = 8*10^7$.

Example 5: Selection of encoded molecules using size-exclusion column

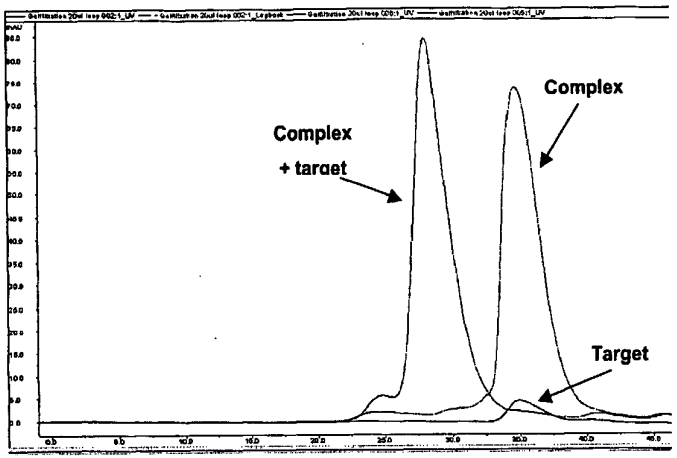
This example illustrates the possibility to use column separation to perform selection on complexes against various targets. In this example, size-exclusion chromatography (SEC) is used, but other types of chromatography can be used where target-bound complexes are separated from the non-bound complexes.

The complex is exemplified in this example by a biotin molecule attached to an oligonucleotide sequence with a predetermined sequence. Thus, the nucleotide sequence of the identifier specifies the identity of the synthetic molecule as biotin. The encoding sequence can have any length and be divided into discrete regions for encoding various building blocks as discussed elsewhere herein. Also, the displayed molecule can have a linear or scaffold structure.

Biotin-AATTCCGGAACATACTAGTCAACATGA

Biotin is known to bind to streptavidin. The binding of biotin to streptavidin will link the identifier to the target molecule and therefore change the identifiers physical and chemical properties, such as e.g. the apparent molecular weight. This change is possible to detect using e.g. size-exclusion chromatography:

78 pmol of the complex molecule was loaded on a Superdex 200, PC 3.2/30 column (ÄKTA-FPLC, AmershamPharmaciaBiotech) and analysed in PBS buffer with a flow rate of 0.050 ml/min. As can be seen below, the complex molecules retention-time was approximately 35 minutes. When the target (83 pmol streptavidin) was analysed under identical conditions the retention-time was approximately the same. The low absorption of the target molecules is due to the wavelength (260 nm) used in the measurement. At this wavelength, the extinction coefficient is high for the nucleotides in the complexes but low for the protein target.



However, when the complex molecules was premixed with the target molecules (78 pmol complex and 83 pmol target incubated for about 1 h in PBS buffer) to allow binding and then analysed under identical conditions, the retention-time change significantly (28 minutes). The change is due to the increase in molecular weight (or hydrodynamic volume) due to the binding of the complex to the target. This will allow the separation of the target-bound complexes from the non-bound complexes. The fraction that contains the complexes and the target molecules are pooled and amplified using appropriate primers. The amplified identifiers can then be used to decode the structures of the enriched displayed molecules.

The strategy of performing column-selection of libraries of bifunctional complexes has two major advantages. First, the enriched (target-bound) complexes are eluted

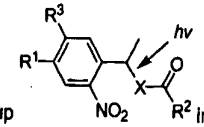
before the non-bound complexes, which will drastically reduce the background from the non-bounded complexes. Secondly, the enrichment on the column will be extensive due to all the separation steps in the pores in the matrix.

The separation of the target-bound complexes using this approach will be dependent on the molecular weight of the complexes but predominantly of the molecular weight of the target. The molecular weight of the target can be adjusted by linking the target to a support that increases the apparent molecular weight. The increased molecular weight will enhance the separation by reducing the retention-time on the column. This can be done using for example a fusion protein, antibody, beads, or cross-linking the target in multimeric form. Thus, the target protein can be expressed as a fusion protein or a specific antibody can be use to increase the molecular weight. The target can be immobilized on small beads that permit separation and the target can be cross-linked using standard reagents to form multimers or cross-linked to a carrier molecule, for example another protein. Preferably, the molecular weight is increase so the target molecules elute in the void volume of the column.

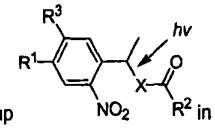
Examples of other types of column separation that can be used are affinity chromatography, hydrophobic interaction chromatography (HIC), and ion-exchange chromatography. Examples of column media, other that Superdex, that can be used in size-exclusion chromatography are: Sephacryl, Sepharose or Sephadex.

Claims

1. A method for identifying a synthetic molecule having affinity towards a target, comprising the steps of
   a) providing a library of bifunctional complexes, wherein each complex of the library comprises a synthetic molecule attached to an identifier, which codes for said molecule,
   b) subjecting, under binding conditions, the library of bifunctional complexes to a target,
   c) removing the non-binding members of the library,
   d) separating the identifiers of complexes comprising synthetic molecules having affinity towards the target, and
   e) decoding the identifiers to establish the identity of the molecule.
2. The method according to claim 1, wherein the synthetic molecule of the library is a non-α-polypeptide.
3. The method of claim 1, wherein the synthetic molecule has a molecular weight less that 2000 Daltons, preferably less than 1000 Dalton, and more preferred less than 500 Daltons.
4. The method according to claims 1 to 3, wherein the identifier uniquely identifies the synthetic molecule.
5. The method according to any of the claim1 to 5, wherein the identifier comprises a sequence of nucleotides.
6. The method according to any of the preceding claims, in which the identifier comprises two or more codons coding for chemical entities which have participated in the synthesis of the synthetic molecule.
7. The method according to claim 6, in which each codon comprises 4 or more nucleotides.
8. The method according to claim 6, wherein the chemical entities are precursors for a structural unit appearing in the synthetic molecule.
9. The method according to claim 6, wherein the chemical entities are not naturally occurring α-amino acids or precursors therefore.
10. The method according to any of the preceding claims, wherein the chemical entities are transferred to the nascent synthetic molecule by a chemical building block, which further comprises an anti-codon.
11. The method according to claim 10, wherein the information of the anti-codon is transferred in conjunction with the chemical entity to the nascent complex.

12. The method according to claims 10 or 11 in which the chemical entities are reacted without enzymatic interaction.
13. The method according to claim 6, wherein the codons are separated by a framing sequence.
14. The method according to any of the preceding claims, wherein the synthesis history of the molecule is established by decoding the identifier.
15. The method according to any of the preceding claims, wherein the synthetic molecule and the identifier are joined by a selectively cleavable linkage.
16. The method according to claim 1, wherein the library comprises two or more different complexes.
17. The method according to claim 1, wherein library comprises 1,000 or more different complexes.
18. The method according to claim 1, wherein the library comprises 1,000,000 or more different complexes.
19. The method according to claim 1, wherein the target is of biological origin.
20. The method according to claims 1 to 19, wherein the target is immobilized on a solid support.
21. The method according to any of the claims 1 to 20, in which a cleavable linkage between the target and the solid support is present.
22. The method according to any of the claims, wherein the separation of identifiers of complexes comprising synthetic molecules having affinity towards the target involves the cleavage of one or more linkages.
23. The method of claims 15 or 21, wherein the linkage is selectively cleavable using electromagnetic radiation, a chemical agent, or an enzyme.
24. The method according to any of the preceding claims, wherein the cleavable linkage comprises a group  in which R^1, R^2, and R^3 independently are the synthetic molecule, the identifier, or a group H or OCH_3, respectively, provided R^2 is either the synthetic molecule or the identifier; and X is selected from the group comprising O, S, or NH.
25. The method according to any of the preceding claims, wherein the removing of the non-binding members of the library in step c) or the separation of step d) includes chromatography.

26. The method of any of the preceding claims, wherein the cleavage of a linkage joining the synthetic molecule and the identifier is preceded by chromatography to separate the identifier.
27. The method according to claims 25 or 26, wherein the chromatography is size-exclusion chromatography.
28. A library of complexes, in which each different complex comprises a synthetic molecule, attached via a cleavable linkage to an identifier which codes for said molecule.
29. A library according to claim 28, wherein the selective cleavable linkage comprises a chemical moiety, which can be cleaved by electromagnetic irradiation.
30. A library according to claim 28 or 29, wherein the cleavable linkage comprises a



group in which R^1, R^2, and R^3 independently are the synthetic molecule, the identifier, or a group H or OCH_3, respectively, provided R^2 is either the synthetic molecule or the identifier; and X is selected from the group comprising O, S, or NH.
31. A library according to any of the claims 28 to 30, wherein the identifier comprises a sequence of nucleotides.
32. A library according to any of the claims 28 to 31, wherein the identifier comprises 2 or more codons, which codes for 2 or more chemical entities incorporated into the synthetic molecule.

1/5

Fig. 1A

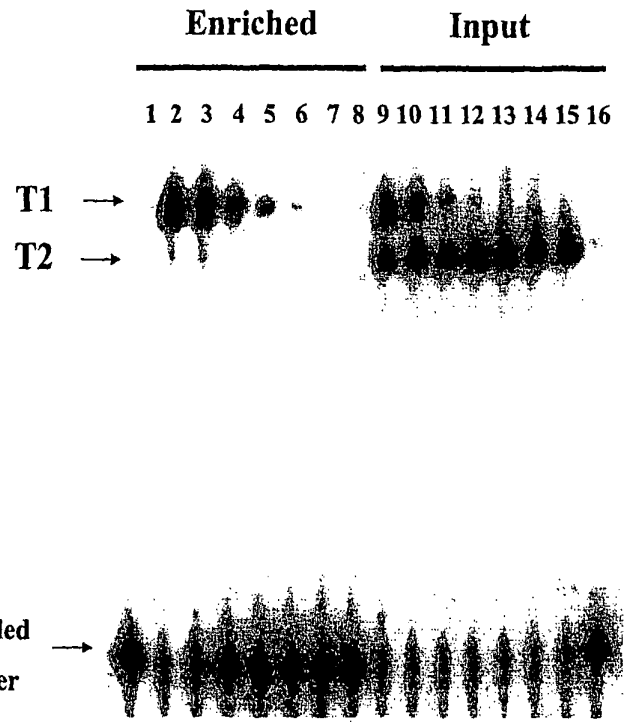


Fig. 1B

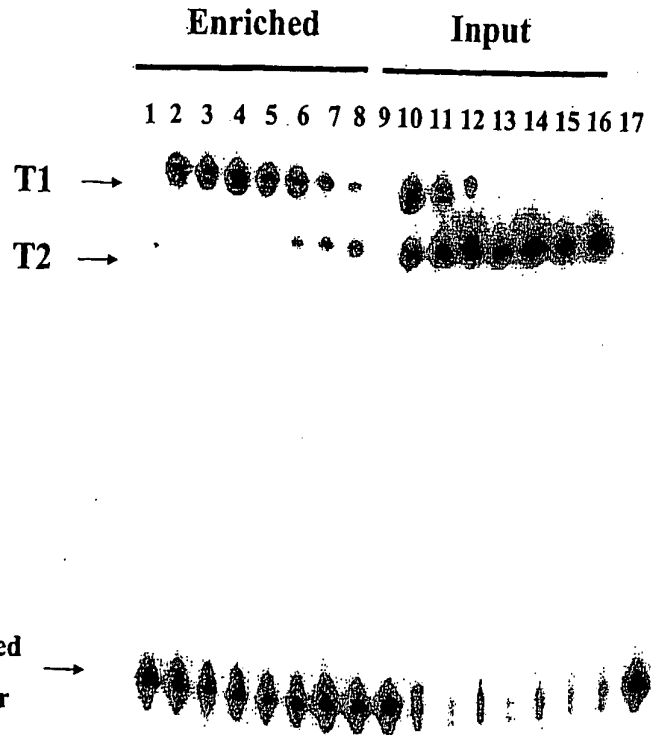


Fig. 2

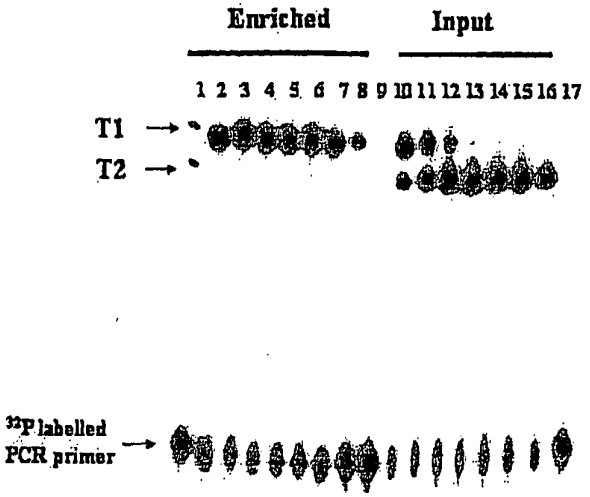


Fig. 3

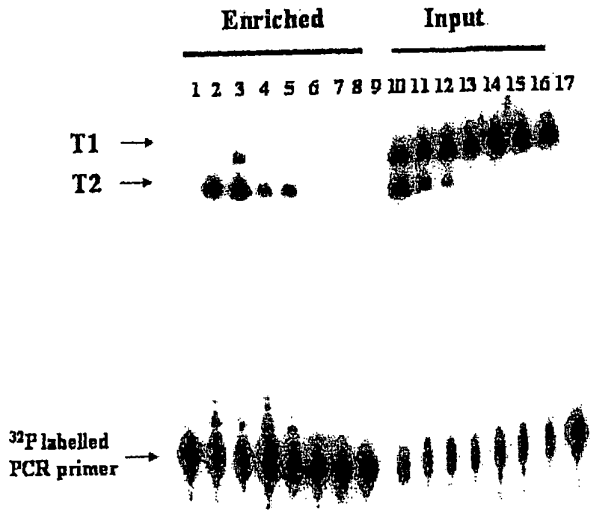


Fig. 4